The comparative statics of multi-prior beliefs and multi-output production*

Paweł Dziewulski[†] John K.-H. Quah[‡]

January 30, 2020

Abstract

Economic decisions often involve maximizing an objective whose value is itself the outcome of another optimization problem. To analyze comparative statics in these models, we introduce a theory of supermodular correspondences. We employ this theory to generalize the notion of first order stochastic dominance to multi-prior beliefs, allowing us to characterize conditions under which greater optimism leads to higher action. We also apply the theory to multi-output production, where it leads to conditions guaranteeing that factors are complements.

Keywords: monotone comparative statics, supermodularity, correspondences, stochastic dominance, multi-output production, maxmin preferences, variational preferences, multiplier preferences

JEL Classification: C61, D21, D24

*We would like to thank Patrick Beissner, Eddie Dekel, Takashi Kamihigashi, Ludvig Sinander, Tomasz Strzalecki, and Bruno Strulovici for their comments and suggestions. The paper was presented at the Canadian Economic Theory Conference (Montreal), Risk, Uncertainty and Decision Conference (Paris), European Workshop in Economic Theory (Berlin), NSF/NBER/CEME Mathematical Economics Conference (Berkeley), the Econometric Society Winter Meeting (Rotterdam), and at seminars in various universities. We would like to thank participants at these gatherings for their comments.

[†]Department of Economics, University of Sussex. E-mail: P.K.Dziewulski@sussex.ac.uk.

[‡]Department of Economics, Johns Hopkins University and Department of Economics, National University of Singapore. E-mail: john.quah@jhu.edu.

1 Introduction

Consider a firm that uses ℓ factors to produce a single good sold at a fixed price. The factors of production are said to be *complements* if a fall in the price of one factor raises the demand for *all* factors (at least weakly). It is well-known that complementarity holds if the production function is supermodular; in this context, supermodularity says that the marginal productivity of a factor is increasing in the level of the other factors.[1] A natural follow up question is to ask what conditions on the production technology will guarantee factor complementarity when the firm is producing multiple output goods. In that case, the firm's production possibility can be represented by a correspondence Γ where set $\Gamma(x)$ consists of all the combinations of output goods that are producible using factors x. Assuming that there are m output goods priced at $q = (q_1, q_2, \ldots, q_m)$, factor complementarity holds if the maximum revenue

$$f(x) \ := \ \max \Big\{ q \cdot y : y \in \Gamma(x) \Big\}$$

is a supermodular function of x.[2] What conditions on Γ will guarantee this?

This issue is one of many in economic modelling that requires supermodularity of a value function after some optimization procedure. For another example, consider an agent who has to take an action under uncertainty. Suppose that the agent's payoff is $g(x, s)$, where $x \in X \subseteq \mathbb{R}$ is the chosen action at state $s \in S \subseteq \mathbb{R}$. The expected utility of action x is therefore $f(x, t) := \int g(x, s) d\lambda(s, t)$, where $t \in T \subseteq \mathbb{R}$ parameterizes the distribution function $\lambda(\cdot, t)$ over S. Suppose that g is such that the marginal payoff of a higher action increases with s, i.e., the function g is supermodular. Then it is reasonable that the *expected* marginal payoff of a higher action should be greater when higher states are more likely. This intuition is correct: if g is a supermodular function of (x, s), then f is supermodular in (x, t) if $\lambda(\cdot, t)$ first order stochastically increases with t. This in turn implies that the optimal action, i.e., argmax $\big\{ f(x, t) : x \in X \big\}$, increases with t.

As a simple application of this result, consider an agent who decides on his savings x in period 1, given uncertainty in his period 2 income, denoted by s. Then

$$g(x, s) \ := \ u(m - x) + \beta u\big(x(1 + r) + s\big), \tag{1}$$

[1] See Topkis (1978), Milgrom and Roberts (1990), and Milgrom and Shannon (1994).

[2] We are assuming here that the firm is a price-taker in all markets. For an alternative interpretation of vector q and correspondence Γ see Section 4.

where u is the per-period utility, β is the discount rate, and r is the interest. In this case, the function g is submodular in (x, s) (equivalently, $g_{xs} \leq 0$) so long as u is concave. Thus, a first order shift in the distribution of period 2 income will *reduce* savings.

Suppose that instead of being an EU-maximizer, the agent is endowed with maxmin preferences as in Gilboa and Schmeidler (1989), so that the ex-ante utility of action x is

$$f(x,t) \;:=\; \min\left\{ \int_S g(x,s)d\lambda(s) : \lambda \in \Lambda(t) \right\}, \tag{2}$$

where $\Lambda(t)$ denotes a *set* of probability distributions over S parameterized by t. Thus, f is the value function arising from Nature choosing $\lambda \in \Lambda(t)$. Assuming that g is supermodular, what conditions on the correspondence Λ will guarantee that f is supermodular (and hence that the optimal choice of x increases with t)? That is, how do we compare *sets of distributions* in a way that generalizes the first order stochastic dominance?

Our results. In Section 2 we formulate different ways of extending the notion of supermodularity to a correspondence $\Gamma : X \to Y$, where X is a lattice and Y is an ordered vector space. Our main results are presented in Section 3. We show that one notion of the supermodularity of Γ is sufficient (and necessary) to guarantee that $f(x) = \max\{\phi(y) : y \in \Gamma(x)\}$ is a supermodular function, for any positive linear functional $\phi : Y \to \mathbb{R}$, while a different but related notion of the supermodularity of Γ characterizes the supermodularity of $f(x) = \min\{\phi(y) : y \in \Gamma(x)\}$. The remainder of the paper is devoted to exploiting these results in different economic contexts.

In Section 4, we apply our main theorems to production analysis. We develop properties on the production correspondence that are sufficient for factors to be complements and provide examples where the property holds. We also explore the conditions under which a change in technology leads to lower (or higher) marginal cost of output.

Section 5 studies the comparative statics of decision-making with maxmin, variational, and multiplier preferences. For each of these models, we formulate what it means for '*beliefs to shift towards higher states.*' For the maxmin model, we show that f, as defined by (2), is supermodular in (x, t) for any supermodular g, if and only if the belief correspondence Λ shifts in the following sense: for any $t' \geq t$ and $\lambda \in \Lambda(t)$, $\lambda' \in \Lambda(t')$, there is $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that

$$\lambda' \succeq \mu,\; \mu' \succeq \lambda, \;\; \text{and} \;\; \tfrac{1}{2}\lambda + \tfrac{1}{2}\lambda' \;=\; \tfrac{1}{2}\mu + \tfrac{1}{2}\mu',$$

where $\succeq$ denotes first order stochastic dominance.[3,4] Returning to our example of the ambiguity averse saver, greater optimism about his period 2 income, captured by a shift in Λ in the sense defined, would lead to lower savings in period 1.

Our definition is a natural way of extending first order stochastic dominance to the comparison of *sets of distributions*; indeed, if $\Lambda(t)$ and $\Lambda(t')$ are singletons then our definition is equivalent to first order stochastic dominance. This definition is also sufficient to guarantee monotone utility comparisons, in the sense that $v(t) = \min \{ \int_S u(s)d\lambda(s) : \lambda \in \Lambda(t)\}$ is increasing in t, for any increasing function u. However, the condition is *not* necessary and in the maxmin model there are at least two useful notions of first order stochastic dominance: one for comparing utilities and another for comparing actions. We provide characterizations for both types of first order stochastic dominance.

We also consider the variational preferences model, where the utility of action x is

$$f(x,t) \;:=\; \min \left\{ \int_S g(x,s)d\lambda(s) + c(\lambda,t) : \lambda \in \triangle_S \right\}.$$

In this context, we characterize the change in the cost function c that captures the notion that the agent considers higher states to be more likely, which leads in turn to a higher marginal utility for higher action. In the case of multiplier preferences, where c is relative entropy, this change in the cost function will occur if the agent's benchmark distribution shifts upwards with respect to the monotone likelihood ratio order.

We consider applications to dynamic programming in Section 6; specifically, we show that the monotone method in Hopenhayn and Prescott (1992) can be extended to the case where, instead of maximizing discounted expected utility, the agent's preference over uncertain utility streams conforms to the maxmin model.

2 Basic concepts

In this section we introduce the basic mathematical concepts that are crucial to our study. A textbook treatment of this material can be found in Topkis (1998).

[3] Note that, if $\lambda' \succeq \lambda$, then μ and μ' can be chosen to be λ and λ', respectively.

[4] This condition holds, for example, if $\Lambda(t) = \{\lambda \in \triangle_S : \nu(\cdot,t) \succeq \lambda \succeq \mu(\cdot,t)\}$, where distributions $\mu(\cdot,t)$ and $\nu(\cdot,t)$ are increasing in t in the first order stochastic sense. Hence, the agent's uncertainty is captured by an interval of distributions, with both its upper and lower bounds increasing in t.

2.1 Orderings, lattices, and comparative statics

A *partial order* $\geq_X$ over a set X is a reflexive, transitive, and antisymmetric binary relation. A partially ordered set, or a *poset*, is a pair $(X, \geq_X)$ consisting of a set X and a partial order $\geq_X$. Whenever it causes no confusion, we denote $(X, \geq_X)$ with X. For any two elements x, x' of a poset X, their *meet*, or the greatest lower bound, is denoted by $x \wedge x'$, and their *join*, or the least upper bound, by $x \vee x'$. A poset X is a *lattice* if for any x, $x' \in X$ both their meet $x \wedge x'$ and their join $x \vee x'$ exist. A subset Y of X is a *sublattice* of X if it contains $y \wedge y'$ and $y \vee y'$ for any y, $y' \in Y$.

A basic example of a lattice is the Euclidean space $\mathbb{R}^\ell$ endowed with the product order $\geq$, i.e., for any vectors x, $x' \in \mathbb{R}^\ell$, we have $x' \geq x$ if $x'_i \geq x_i$, for all $i = 1, \ldots, \ell$. In this case, the meet $x \wedge x'$ and the join $x \vee x'$ are given by $(x \wedge x')_i = \min\{x_i, x'_i\}$ and $(x \vee x')_i = \max\{x_i, x'_i\}$, for all $i = 1, \ldots, \ell$. For the purposes of this paper, the second important example of a lattice is $(\triangle_S, \succeq)$, where $\triangle_S$ is the set of cumulative distribution functions defined on $S \subseteq \mathbb{R}$ and $\lambda' \succeq \lambda$ if λ' first order stochastically dominates λ (which means that $\lambda'(s) \leq \lambda(s)$, for all $s \in S$). Then, for two distribution λ and μ, the join $\lambda \vee \mu$ is given by $(\lambda \vee \mu)(s) = \min\{\lambda(s), \mu(s)\}$ and the meet $\lambda \wedge \mu$ by $(\lambda \wedge \mu)(s) = \max\{\lambda(s), \mu(s)\}$.

Our main theoretical results are applicable to correspondences that map a lattice to an ordered vector space. An *ordered vector space* $(Y, \geq_Y)$ is a real vector space Y endowed with a partial order $\geq_Y$ that preserves the vector space operations, i.e., for any y, $y' \in Y$, if $y' \geq_Y y$ then $(y' + z) \geq_Y (y + z)$ and $\alpha y' \geq_Y \alpha y$, for any $z \in Y$ and $\alpha \geq 0$. For the purposes of the applications in this paper, the relevant ordered vector space is simply the Euclidean space endowed with the product order.

Lattice-theoretic concepts play an important role in the study of comparative statics. For any two subsets Y, Y' of a lattice X, we say that Y' dominates Y in the *strong set order induced by* $\geq_X$, if for any $y \in Y$ and $y' \in Y'$, we have $y \wedge y' \in Y$ and $y \vee y' \in Y'$. Whenever Y' and Y contain their greatest elements y' and y, respectively, then Y' dominates Y in the strong set order only if $y' \geq_X y$.[5] While the strong set order is not complete, it is transitive over subsets of X (see Topkis, 1978).

A function $f : X \to \mathbb{R}$ defined over a lattice X is *supermodular* if for any elements x,

[5] Similarly, if Y' and Y contain their least elements y' and y respectively, then Y' dominates Y in the strong set order only if $y' \geq_X y$. Moreover, whenever $Y = \{y\}$ and $Y' = \{y'\}$ (i.e., the sets are singletons) then $y' \geq_X y$ if and only if Y' dominates Y in the strong set order.

$x' \in X$, we have $f(x \wedge x') + f(x \vee x') \geq f(x) + f(x')$. We say that f is *submodular* if $(-f)$ is supermodular.

Let X be a lattice and T be a partially ordered set. A function $f : X \times T \to \mathbb{R}$ is said to have *increasing differences* if, for all $x' \geq_X x$, the difference $\delta(t) = f(x', t) - f(x, t)$ is increasing in t. This notion is closely related to supermodularity; indeed, if $T \subseteq \mathbb{R}$, then it is straightforward to check that the function $f(x, t)$ is supermodular in (x, t) with respect to the product order on $X \times T$ if and only if it is supermodular in x and has increasing differences in (x, t).

Topkis (1978) shows that whenever the function $f : X \times T \to \mathbb{R}$ is supermodular in x, then the set of maximizers $\Phi(t) = \arg\max \{f(x, t) : x \in X\}$ is a sublattice of X. If, in addition, f has increasing differences in (x, t), then this set increases in t with respect to the strong set order, i.e., $\Phi(t')$ dominates $\Phi(t)$ in the strong set order whenever $t' \geq t$. We shall refer to this result as the *Monotone Comparative Statics* (MCS) theorem.[6] It is also known that if $\Phi(t)$ is a *compact* sublattice of a Euclidean space, then it must contain the least and the greatest elements, and both will be increasing in t.

2.2 Upper and lower supermodular correspondences

Suppose that $(X, \geq_X)$ is a lattice and $(Y, \geq_Y)$ is an ordered vector space. A correspondence $\Gamma : X \to Y$ is *upper supermodular* if for any two elements x, $x' \in X$ and $y \in \Gamma(x)$, $y' \in \Gamma(x')$, there is $z \in \mathrm{co}\Gamma(x \wedge x')$ and $z' \in \mathrm{co}\Gamma(x \vee x')$ such that

$$z + z' \ \geq_Y \ y + y'.\text{[7]} \tag{3}$$

Equivalently, the condition can be expressed in terms of average vectors that satisfy $(1/2)z + (1/2)z' \geq_Y (1/2)y + (1/2)y'$. See Figure 1 for a graphical interpretation.

The correspondence Γ is *lower supermodular* if for any x, $x' \in X$ and $z \in \Gamma(x \wedge x')$, $z' \in \Gamma(x \vee x')$ there are vectors $y \in \mathrm{co}\Gamma(x)$, $y' \in \mathrm{co}\Gamma(x')$ that satisfy (3).[8] Finally, the

[6] Milgrom and Shannon (1994) present a well-known generalization of the MCS theorem where supermodularity and increasing differences are replaced by their ordinal counterparts.

[7] By co$A$ we denote the convex hull of set A.

[8] Notice that, the distinction between upper and lower supermodularity disappears if Γ is a function, i.e., Γ is a singleton-valued, rather than a set-valued correspondence. It is easy to construct correspondences that are upper supermodular but not lower supermodular, or vice versa. For example, let $X = \{x, x', (x \vee x'), (x \wedge x')\}$, where x and x' are unordered points in $\mathbb{R}^2$ and define $\Gamma : X \to \mathbb{R}$ by $\Gamma(x) = \Gamma(x') = \{1\}$, and $\Gamma(x \wedge x') = \Gamma(x \vee x') = [0.5, 1]$. Then Γ is upper supermodular since 1 is in

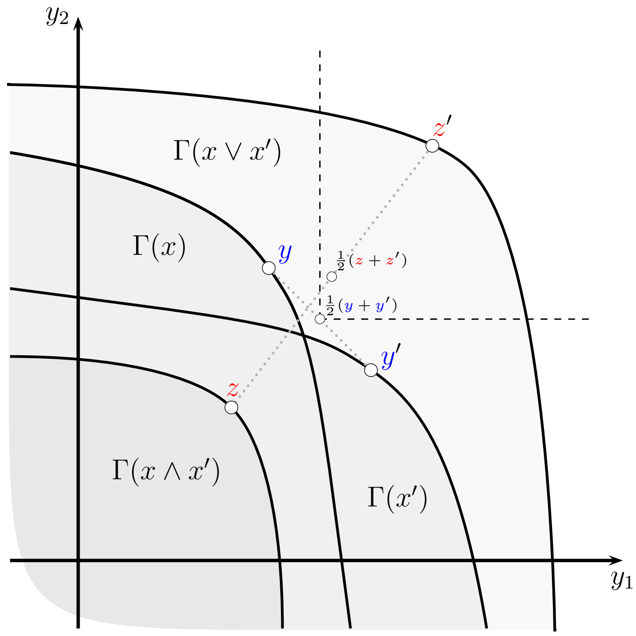

Figure 1: An upper supermodular correspondence $\Gamma : X \to Y = \mathbb{R}^2$.

correspondence is *supermodular* if it is both upper and lower supermodular.

Submodularity of correspondences can be defined analogously. The correspondence Γ is *upper submodular* if for any x, $x' \in X$ and $y \in \Gamma(x)$, $y' \in \Gamma(x')$, there is some $z \in \text{co}\Gamma(x \wedge x')$, $z' \in \text{co}\Gamma(x \vee x')$ that satisfy

$$y + y' \;\; \geq_Y \;\; z + z';$$

equivalently, this means that $(-\Gamma)$ is an upper supermodular correspondence. One may define lower submodularity and submodularity in an analogous fashion.

Our definition of supermodular correspondences generalizes the familiar notion of supermodularity applied to real-valued functions, presented at the beginning of this section. It also extends the concept of *stochastic supermodularity* introduced in Topkis (1968) to correspondences;[9] a function mapping a lattice to the set of probability measures on some measurable space is said to be stochastically supermodular if condition (3) holds with $\geq_Y$ representing the first order stochastic dominance.

both $\Gamma(x \vee x')$ and $\Gamma(x \wedge x')$. However, it is not lower supermodular. Indeed, choose $z = z' = 0.5$, where $z \in \Gamma(x \wedge x')$ and $z' \in \Gamma(x \vee x')$; then $y + y' = 2 > 1 = z + z'$, for all $y \in \Gamma(x)$ and $y' \in \Gamma(x')$.

[9] Although Topkis (1968) refers to this property as *stochastic convexity*, the term *stochastic supermodularity* is also commonly used (see Curtat, 1996 or Balbus, Reffett, and Woźny, 2014).

Suppose that $\Gamma : X \to Y$ has convex and *downward comprehensive* values, where latter means that $y \in \Gamma(x)$ and $y \geq_Y z$ implies $z \in \Gamma(x)$. Then the correspondence Γ is upper supermodular if and only if, for all x, $x' \in X$,

$$\Gamma(x \wedge x') + \Gamma(x \vee x') \ \supseteq \ \Gamma(x) + \Gamma(x'). \tag{4}$$

The fact that (4) implies upper supermodularity is clear and does not even require Γ to have downward comprehensive values. To show the converse, suppose that Γ is upper supermodular. Thus, for any $y \in \Gamma(x)$, $y' \in \Gamma(x')$ there is some $z \in \Gamma(x \wedge x')$, $z' \in \Gamma(x \vee x')$ such that $z + z' \geq_Y y + y'$. In particular, we have $z \geq_Y (y + y' - z')$. Since Γ is downward comprehensive, it must be that $(y + y' - z') \in \Gamma(x \wedge x')$. Consequently, this implies that $(y + y') = (y + y' - z') + z'$ belongs to $\Gamma(x \wedge x') + \Gamma(x \vee x')$.

A special case of property (4) appears in the study of cooperative games with non-transferable utility. In that context, X is the collection of coalitions of a finite set N of players in a game, i.e., the power set of N; when endowed with the set inclusion order $\supseteq$, the pair $(X, \supseteq)$ forms a lattice. For any coalition x, set $\Gamma(x) \subseteq \mathbb{R}^N$ consists of utility profiles (across all players in the game) that could result from the formation of that coalition. The game is said to be *cardinally convex* if (4) holds (see Sharkey, 1981, Section 2); in other words, whenever the correspondence Γ is upper supermodular.

2.3 Generating supermodular correspondences

In this subsection we list some simple results on generating or preserving the supermodular property on correspondences.

Fact 1. Upper supermodularity is preserved by downward comprehensive extensions. To be precise, whenever the correspondence $\Gamma : X \to Y$ is upper supermodular, then so is $\bar{\Gamma}(x) = \{y \in Y : y \leq_Y z, \text{ for some } z \in \Gamma(x)\}$. Similarly, lower supermodularity is preserved by upward comprehensive extensions.

Fact 2. Upper and lower supermodularity are preserved by summation, i.e., for any upper (lower) supermodular correspondences Γ, $\Lambda : X \to Y$, the correspondence $\Omega(x) = \alpha\Gamma(x) + \beta\Lambda(x)$ is an upper (lower) supermodular, for any positive scalars $\alpha, \beta \geq 0$.

Fact 3. The functions $g_i : X \to \mathbb{R}$ are supermodular over a lattice X, for all $i = 1, \ldots, \ell$, if and only if the map $G : X \to \mathbb{R}^\ell$, given by $G(x) = \big(g_1(x), \ldots, g_\ell(x)\big)$, is a supermodular function, i.e., we have $G(x \wedge x') + G(x \vee x') \geq G(x) + G(x')$, for all x, $x' \in X$, where $\geq$ denotes the natural product order on $\mathbb{R}^\ell$.

Fact 4. Let $\Gamma_i : X_i \to Y$ be a correspondence from $X_i \subseteq \mathbb{R}$ to an ordered vector space Y, for $i = 1, 2$. Then the map $\Lambda : X_1 \times X_2 \to Y$, where $\Lambda(x_1, x_2) = \Gamma_1(x_1) + \Gamma_2(x_2)$, is a supermodular correspondence (in fact, it is also submodular).

Fact 5. For an arbitrary subset Z of an ordered vector space Y, a supermodular function $h : X \to Y$ over a lattice X, and positive scalars α and β, the mapping $\Gamma : X \to Y$, given by $\Gamma(x) = \big\{\alpha y + \beta h(x) : y \in Z\big\}$, is a supermodular correspondence.

Fact 6. Let Z be a subset of an ordered vector space Y such that $z \geq 0$, for all $z \in Z$. For any positive, supermodular function $h : X \to \mathbb{R}_+$ over a lattice X, the correspondence $\Gamma : X \to Y$ given by $\Gamma(x) = \big\{h(x)z : z \in Z\big\}$ is supermodular.

This claim requires a short proof. Lemma 5.27 in Aliprantis and Border (2006) guarantees that $\alpha \text{co}Z + \beta \text{co}Z = (\alpha + \beta)\text{co}Z$, for any positive scalars α and β. To show that the map Γ is upper supermodular, take any $h(x)y \in \Gamma(x)$ and $h(x')y' \in \Gamma(x')$. Given the above property of set Z, there is a vector $v \in \text{co}Z$ such that $h(x)y + h(x')y' = \big[h(x) + h(x')\big]v$. Since h is supermodular and Z is nonnegative,

$$\big[h(x) + h(x')\big]v \;\leq\; \big[h(x \wedge x') + h(x \vee x')\big]v.$$

Since $h(x \wedge x')v \in \text{co}\Gamma(x \wedge x')$ and $h(x \vee x')v \in \text{co}\Gamma(x \vee x')$, this concludes the proof. An analogous argument guarantees that Γ is also lower supermodular.

Fact 7. Let X, T be lattices and Z be a sublattice of $X \times T$ (endowed with the product order). By X_Z we denote the set of elements in X for which there is some $t \in T$ such that $(x, t) \in Z$; it is straightforward to check that X_Z is a sublattice of X. Suppose that $h : Z \to Y$ is a supermodular function, where Y is an ordered real vector space. Then the correspondence $\Gamma : X_Z \to Y$, given by

$$\Gamma(x) \;:=\; \Big\{h(x, t) : (x, t) \in Z\Big\},$$

is upper supermodular. Indeed, take any $y \in \Gamma(x)$ and $y' \in \Gamma(x')$. By the definition of Γ, there is some t and t' in T such that $y = h(x, t)$ and $y' = h(x', t')$. Moreover, the supermodularity of function h implies that

$$h\big((x \wedge x'), (t \wedge t')\big) + h\big((x \vee x'), (t \vee t')\big) \;\geq\; h(x,t) + h(x',t'),$$

where $h\big((x,t) \wedge (x',t')\big)$ belongs to $\Gamma(x \wedge x')$ and $h\big((x,t) \vee (x',t')\big)$ to $\Gamma(x \vee x')$.

3 Value functions of supermodular correspondences

In this section we present our main theorems on supermodular correspondences. While the proofs are simple, these results lead naturally to a wide range of applications.

Main Theorem. *Suppose that X is a lattice and Y is an ordered vector space. For any positive linear functional $\phi : Y \to \mathbb{R}$,*[10]

(i) *if correspondence $\Gamma : X \to Y$ is upper supermodular then the function $f : X \to \mathbb{R}$, given by $f(x) := \max\big\{\phi(y) : y \in \Gamma(x)\big\}$, is supermodular;*[11]

(ii) *if correspondence $\Gamma : X \to Y$ is lower supermodular then the function $f : X \to \mathbb{R}$, given by $f(x) := \min\big\{\phi(y) : y \in \Gamma(x)\big\}$, is supermodular.*

Proof. To show (i), take any x, $x' \in X$ and $y \in \Gamma(x)$, $y' \in \Gamma(x')$. By the upper supermodularity of Γ, there is $z \in \mathrm{co}\Gamma(x \wedge x')$ and $z' \in \mathrm{co}\Gamma(x \vee x')$ such that $z + z' \geq_Y y + y'$. Therefore, for any positive linear functional $\phi : Y \to \mathbb{R}$,

$$\begin{aligned} \phi(y) + \phi(y') \;&=\; \phi\big(y + y'\big) \;\leq\; \phi\big(z + z'\big) \;=\; \phi(z) + \phi(z') \\ &\leq\; \max\big\{\phi(v) : v \in \Gamma(x \wedge x')\big\} + \max\big\{\phi(v) : v \in \Gamma(x \vee x')\big\} \\ &=\; f(x \wedge x') + f(x \vee x'), \end{aligned}$$

where the first inequality follows from ϕ being positive and the second is implied by the definition of maximum and the fact that $\max\big\{\phi(y) : y \in A\big\} = \max\big\{\phi(y) : y \in \mathrm{co}A\big\}$ (for any set $A \subseteq Y$). Taking the maximum over the left-hand side of the inequality, we conclude that $f(x) + f(x') \leq f(x \wedge x') + f(x \vee x')$. Hence, f is supermodular.

[10] A linear functional $\phi : Y \to \mathbb{R}$ is *positive*, whenever $y \geq_Y z$ implies $\phi(y) \geq \phi(z)$, for all y, z in Y.

[11] We shall assume throughout this paper that a solution exists to any optimization problem we consider, so that we could always speak of the maximum (minimum) rather than the supremum (infimum). That said, it is easy to check that both the Main Theorem and Main Theorem $(*)$ remain valid if the existence of an optimum is not guaranteed and we have to replace max (min) with sup (inf).

To prove (ii), take any $z \in \Gamma(x \wedge x')$, $z' \in \Gamma(x \vee x')$. By the lower supermodularity of Γ, there is some $y \in \Gamma(x)$, $y' \in \Gamma(x')$ such that $z + z' \geq_Y y + y'$. Therefore,

$$\begin{aligned}\phi(z) + \phi(z') \;=\; \phi(z+z') \;\geq\; \phi(y+y') \;=\; \phi(y) + \phi(y') \\ \geq\; \min\left\{\phi(v) : v \in \Gamma(x')\right\} + \min\left\{\phi(v) : v \in \Gamma(x)\right\} \;=\; f(x) + f(x'),\end{aligned}$$

where the first inequality follows from ϕ being positive and the second is implied by the definition of minimum and the fact that $\min\left\{\phi(y) : y \in A\right\} = \min\left\{\phi(y) : y \in \mathrm{co}A\right\}$ (for any $A \subseteq Y$). By taking the minimum over the left-hand side of this inequality, we obtain $f(x \wedge x') + f(x \vee x') \geq f(x) + f(x')$. **QED**

In some applications one would like to guarantee the submodular property of the value function. In those instances the following analogue to the Main Theorem applies; we skip the proof since it is similar.

Main Theorem $(*)$**.** *Suppose that X is a lattice and Y is an ordered vector space. For any positive linear functional $\phi : Y \to \mathbb{R}$,*

(i) *if correspondence $\Gamma : X \to Y$ is upper submodular then function $f : X \to \mathbb{R}$, given by $f(x) := \min\left\{\phi(y) : y \in \Gamma(x)\right\}$, is submodular;*

(ii) *if correspondence $\Gamma : X \to Y$ is lower submodular then function $f : X \to \mathbb{R}$, given by $f(x) := \max\left\{\phi(y) : y \in \Gamma(x)\right\}$, is submodular.*

The assumptions in the Main Theorem are essentially tight. The following result gives a converse to the theorem in the case where Y is a Euclidean space.[12]

Proposition 1. *Suppose that X is a lattice, Y is a Euclidean space, and the correspondence $\Gamma : X \to Y$ is such that set $\mathrm{co}\,\Gamma(x) + \mathrm{co}\,\Gamma(x')$ is closed for any $x, x' \in X$.*

(i) *If the function $f : X \to \mathbb{R}$, given by $f(x) := \max\left\{\phi(y) : y \in \Gamma(x)\right\}$, is supermodular for any positive linear functional $\phi : Y \to \mathbb{R}$, then Γ is upper supermodular.*

(ii) *If the function $f : X \to \mathbb{R}$, given by $f(x) := \min\left\{\phi(y) : y \in \Gamma(x)\right\}$, is supermodular for any positive linear functional $\phi : Y \to \mathbb{R}$, then Γ is lower supermodular.*

[12] We omit the converse to Main Theorem $(*)$, which has a proof similar to that for Proposition 1.

Remark. The above proposition requires that $\operatorname{co}\Gamma(x) + \operatorname{co}\Gamma(x')$ be a closed set, for all $x, x' \in X$. The assumption holds naturally in the economic applications discussed in Sections 4–6. In particular, it holds when Γ is compact-valued; more generally, it holds when, for all x, $\Gamma(x)$ is closed and bounded from below or closed and bounded from above. This follows from Proposition 2.38 in Border (1985).[13]

The proof of Proposition 1, which can be found in the Appendix, proceeds by contradiction. We show that whenever a correspondence is not upper supermodular, we can apply the separating hyperplane theorem to produce a positive linear functional ϕ for which the maximal value f is not a supermodular function.

4 Applications to production analysis

The results developed in the last section lead to many applications. We begin our discussion with the most obvious of these, which is the application to multi-output production.

4.1 Complementarity in multi-output production

A firm is endowed with a technology that employs ℓ inputs to manufacture m output goods. Following McFadden (1966, 1978) and Jacobsen (1970) we represent its production possibilities with a production correspondence $\Gamma : \mathbb{R}^{\ell}_{+} \to \mathbb{R}^{m}_{+}$ that maps input vector $x \in \mathbb{R}^{\ell}_{+}$ to $\Gamma(x)$, the set of output vectors that are feasible given the use of x.

Conditional on strictly positive prices for inputs, $p \in \mathbb{R}^{\ell}_{++}$, and for outputs, $q \in \mathbb{R}^{m}_{++}$, the problem of the firm is to choose input x to maximise

$$\pi(x, p) \;:=\; \max\left\{q \cdot y : y \in \Gamma(x)\right\} - p \cdot x.$$

Even though we interpret vectors y as output profiles, there is a related but slightly different interpretation. Suppose that the firm is operating in a risky environment with m states of the world. Then, vector y determines all the contingent revenues that the firm may choose, when the input vector x is employed. If q is the probability distribution over different states, then $q \cdot y$ is the expected revenue under profile y.

[13] A subset $A \subseteq \mathbb{R}^{\ell}$ is *bounded from below* if there is some $z \in \mathbb{R}^{\ell}$ such that $z \leq y$, for all $y \in A$ and it is bounded from *above* if there is some $z \in \mathbb{R}^{\ell}$ such that $z \geq y$, for all $y \in A$.

We are interested in conditions on Γ guaranteeing that all inputs are *complements* in the sense that the demand for *all* inputs increase when the price of one input drops; in formal terms, complementarity requires that the set of optimal input vectors, $\Phi(p) = \arg\max\left\{\pi(x,p) : x \in X\right\}$, decreases in p with respect to the strong set order.

Proposition 2. *Inputs are complements if Γ is upper supermodular.*

This proposition follows from the Main Theorem and the MCS theorem. First, for each $x \in X$, the firm determines the maximal revenue that is achievable, which is $f(x) := \max\{q \cdot y : y \in \Gamma(x)\}$. In the second step, the firm chooses $x \in \mathbb{R}^\ell_+$ to maximise profit $\pi(x,p) = f(x) - p \cdot x$. From this observation and the MCS theorem, we know that inputs are complements if function π is supermodular in x and has increasing differences in $(x, -p)$. The latter is always true, while the former is satisfied when f is supermodular. The Main Theorem guarantees that f is supermodular if the production correspondence Γ is upper supermodular. Furthermore, we know from Proposition 1 that upper supermodularity is necessary for the supermodularity of f provided Γ has convex and closed values that are bounded from above.

The following examples are applications of Proposition 2.

Example 1. Consider a production technology with three inputs and two outputs (or state contingent revenues), where

$$\Gamma(x_1,x_2,x_3) \;:=\; \left\{(y_1,y_2) \in \mathbb{R}^2 : y_1 \le \sqrt[3]{x_1\, x_2\, t},\; y_2 \le \sqrt{x_1} + \sqrt{x_3 - t},\; \text{for } t \in [0, x_3]\right\}.$$

In this case, input 1 is non-rivalrous since it can be used in its entirety to produce both outputs. On the other hand, input 3 has to be shared between the two productions, while input 2 is only used in the production of good 1. We claim that this correspondence is upper supermodular. Indeed, notice that the set

$$Z \;:=\; \left\{(x_1,x_2,x_3,t) \in \mathbb{R}^4 : x_i \ge 0, \text{ for } i = 1,2,3, \text{ and } t \in [0,x_3]\right\}$$

is a sublattice of $\mathbb{R}^4$. Moreover, $h : Z \to \mathbb{R}^2$, where $h(x,t) := \left(\sqrt[3]{x_1\, x_2\, t}, \sqrt{x_1} + \sqrt{x_3 - t}\right)$, is a supermodular function. Therefore, by Fact 7, $\tilde{\Gamma}(x) := \{h(x,t) : (x,t) \in Z\}$ is an upper supermodular correspondence. Given that Γ is the downward comprehensive extension of the mapping $\tilde{\Gamma}$, Fact 1 implies that Γ is also upper supermodular.

Example 2. Suppose that

$$\Gamma(x) \;:=\; \Big\{y \in \mathbb{R}^m_+ : g(x) \geq h(y)\Big\},$$

where $g : \mathbb{R}^\ell_+ \to \mathbb{R}_+$ and $h : \mathbb{R}^m_+ \to \mathbb{R}_+$ are strictly increasing functions. (We could interpret $g(x)$ as the level of some intermediate good which can be produced with x; this intermediate good is then transformed into different output goods via the function h.) We claim that, whenever g is supermodular and h is homogeneous of degree 1, the correspondence Γ is a supermodular, and thus upper supermodular, so that Proposition 2 applies. Indeed, we can write $\Gamma(x) = g(x)Z$, where $Z = \big\{z \in \mathbb{R}^\ell_+ : 1 \geq h(z)\big\}$, since h is homogeneous of degree 1; Fact 6 guarantees that Γ is supermodular.

4.2 Technological change and marginal cost

In this application, we consider a firm that produces a *single* good using ℓ inputs and investigate the conditions under which technological change reduces the marginal cost of production.

Let $\big\{F(t,\cdot)\big\}_{t\in T}$ be a family of production functions. At input prices $p \in \mathbb{R}^\ell_{++}$, the production function $F(t,\cdot) : \mathbb{R}^\ell_+ \to \mathbb{R}$ induces the cost function $C(t,\cdot) : \mathbb{R}_+ \to \mathbb{R}$, where

$$C(t,q) \;:=\; \min\Big\{p \cdot y : F(t,y) \geq q\Big\}.$$

We assume that output q generates revenue $B(q)$ and the firm's objective is to choose q to maximize profit, which is $B(q) - C(t,q)$. Irrespective of the precise shape of B, we know that the profit-maximizing output increases with t whenever C is a submodular function, since this guarantees that the profit function is supermodular in (t,q).[14] Notice that the submodularity of C is just another way of saying that the marginal cost of output, $C_q(t,q)$, is decreasing in t, for all q.

So when does marginal cost decrease with t? One may be tempted to think that that occurs whenever the technology change raises output (formally, if $F(t,y)$ is increasing in t at each y) but that is *not* the case.[15] By the Main Theorem $(*)$, the function $C(t,q)$ is

[14] Furthermore, it is straightforward to check that the submodularity of C is also necessary for this to hold if the benefit function B is allowed to take different shapes.

[15] For example, suppose $F(t,y) = y^\beta + t$, where y is the level of the unique input. Obviously, $F(t,y)$ increases with t, but $C(t,q) = (q-t)^{1/\beta}$ and marginal cost $C_q(t,q)$ falls with t if and only if $\beta \leq 1$.

submodular in (q,t) if the correspondence $\Gamma : T \times \mathbb{R}_+ \to \mathbb{R}^\ell_+$, given by

$$\Gamma(t,q) \; := \; \Big\{ y \in \mathbb{R}^\ell_+ : F(t,y) \geq q \Big\}, \tag{5}$$

is upper submodular. Obviously, $\Gamma(t,q)$ is bounded from below by 0, and it will have closed values provided F is continuous; under this condition, the upper submodularity of Γ is also *necessary* for the submodularity of C (see remark following Proposition 1).

Example 3. Suppose $F(t,y) = t\,G(y)$ and G is homogeneous of degree 1. That is, the function $F(t,\cdot)$ exhibits constant returns to scale, but need not be quasiconcave. In this case, there is no need to appeal to our result — the marginal cost is constant at all output levels; if it is c^* when $t = 1$ then then it will be c^*/t for any other $t > 0$, so marginal cost falls with t. Nonetheless it is instructive to see how it fits within our framework. It is straightforward to check that, given constant returns to scale, $\Gamma(t,q) = (q/t)Z$, where $Z = \big\{ y : G(y) \geq 1 \big\}$. This set is upper submodular, by Fact 6.

Example 4. Suppose that $F : T \times \mathbb{R}^\ell_+ \to \mathbb{R}_+$ is increasing and supermodular over $T \times \mathbb{R}^\ell_+$, and $F(t,\cdot)$ is continuous and concave in $y \in \mathbb{R}^\ell_+$, for each $t \in T$. We show in the Appendix that this suffices for Γ to be upper submodular, so that C is a submodular function. In particular, if $F(t,y) = t\,G(y)$, where $G : \mathbb{R}^\ell_+ \to \mathbb{R}_+$ is a continuous, increasing, concave, and supermodular function, then we know that the marginal costs related to the production functions $F(t,\cdot)$ decrease with $t \in T \subseteq \mathbb{R}_+$.

5 Supermodular correspondences and uncertainty

In this section, we consider an agent who has to pick an action under uncertainty. Suppose that the possible states of the world are represented by a set $S \subseteq \mathbb{R}$; to keep our exposition focused on the essentials we assume that $S = \{s_1, s_2, \ldots, s_\ell, s_{\ell+1}\}$ is finite, where $s_1 < s_2 < \ldots < s_\ell < s_{\ell+1}$. We denote the distributions on S by $\triangle_S$. As we had pointed out in Section 2.1, $(\triangle_S, \succeq)$ forms a lattice, where $\lambda \succeq \mu$ if λ first order stochastically dominates μ; this feature plays an important role in our analysis.[16]

[16] For this property S need not be finite, but it is crucial that S is a subset of $\mathbb{R}$. Although first order stochastic dominance can be naturally extended to distributions over multi-dimensional spaces, in such a case $(\triangle_S, \succeq)$ would no longer constitute a lattice.

First order stochastic dominance is a concept of fundamental importance because it allows us to *compare distributions by expected utility*: $\lambda \succeq \mu$ if and only if $\int_S u(s)d\lambda(s) \geq \int_S u(s)d\mu(s)$ for all increasing functions $u : S \to \mathbb{R}$. Furthermore, this basic result has a simple and widely-used corollary that also allows us to *compare the actions* of an agent maximizing expected utility. To be specific, consider an agent who chooses an action from a set $X \subseteq \mathbb{R}$. The agent's utility from choosing action x is $g(x, s)$ whenever state s is realized. Let $\lambda(\cdot, t)$ be a distribution over S (parameterized by $t \in T \subseteq \mathbb{R}$) which captures the agent's belief about the likelihood of different states. Then, the expected utility of taking action x is $f(x, t) = \int_S g(x, s)d\lambda(s, t)$. Now suppose that g is supermodular and λ is ordered by first order stochastic dominance in the sense that $\lambda(\cdot, t') \succeq \lambda(\cdot, t)$ whenever $t' \geq t$. In such a case, if $x' \geq x$, then

$$f(x', t) - f(x, t) \;=\; \int_S \big[g(x', s) - g(x, s)\big]\, d\lambda(s, t),$$

will be increasing in t since $\delta(s) = g(x', s) - g(x, s)$ is increasing in s. In other words, f is supermodular in (x, t), which (by the MCS theorem) guarantees that the correspondence $\Phi(t) = \text{argmax}\,\big\{f(x, t) : x \in X\big\}$ increases with t in the strong set order.

Our objective in this section is to extend this simple result on comparative statics to some widely-used multi-prior models of decision-making under uncertainty.

5.1 First order stochastic dominance in the maxmin model

In the *maxmin* model of Gilboa and Schmeidler (1989), the agent evaluates an uncertain environment not with a single distribution over the possible states of the world but with a set of distributions $\Lambda \subseteq \triangle_S$. If $u(s)$ is the utility when s is realized, then the agent's utility in this uncertain environment is

$$\min \left\{ \int_S u(s)d\lambda(s) : \lambda \in \Lambda \right\}.$$

We know that, when Λ consists of just one distribution, a first order stochastic shift in the distribution will lead to higher utility, assuming that u is increasing in s. This leads naturally to the following question: what shift in the *set* of beliefs would guarantee that there is an increase in utility? The following proposition provides the precise answer.

Proposition 3. *Suppose the correspondence $\Lambda : T \to \triangle_S$ has compact and convex values. Then the following statements are equivalent.*

(i) *Correspondence* Λ *satisfies the following property:*

(F1) *if* $t' \geq t$*, then for any* $\lambda' \in \Lambda(t')$ *there is* $\lambda \in \Lambda(t)$ *such that* $\lambda' \succeq \lambda$*.*

(ii) *For any increasing function* $u : S \to \mathbb{R}$*, the function* $v : T \to \mathbb{R}$*, given by* $v(t) := \min \left\{ \int_S u(s) d\lambda(s) : \lambda \in \Lambda(t) \right\}$ *is increasing in* t*.*

This proposition gives us one natural way of defining first order stochastic dominance between *sets* of distributions since (F1) characterizes higher ex ante utility when u is increasing. The sufficiency of (F1) is easy to show. Indeed, take any $t' \geq t$ and $\lambda' \in \Lambda(t')$. By (F1), there is $\lambda \in \Lambda(t)$ such that $\lambda' \succeq \lambda$. Thus, for any increasing u,

$$\int_S u(s) d\lambda'(s) \;\geq\; \int_S u(s) d\lambda(s) \;\geq\; \min \left\{ \int_S u(s) d\nu(s) : \nu \in \Lambda(t) \right\}.$$

Taking the minimum over the left-hand side, we obtain $v(t') \geq v(t)$. The proof of the necessity of (F1) is found in the Appendix.

A natural follow-up question is whether (F1) is also sufficient to guarantee monotone comparative statics. More precisely, let the utility from choosing action x be $g(x, s)$ when state s is realized and suppose that g is a supermodular function. Then one could ask if (F1) guarantees that the function

$$f(x, t) \;:=\; \min \left\{ \int_S g(x, s) d\lambda(s) : \lambda \in \Lambda(t) \right\} \tag{6}$$

is supermodular in (x, t). The following example shows that this is not true.

Example 5. Suppose that $X = \{0, 1\}$ and $S = \{s_1, s_2, s_3\}$. The distribution λ is given by $\lambda(s_1) = 1/2$ and $\lambda(s_2) = 3/4$. The distribution λ' is given by $\lambda'(s_1) = \lambda'(s_2) = 1/2$ and $\mu(s_1) = 1/4$, $\mu(s_2) = 7/8$. Suppose that $T = \{t, t'\}$, where $t' > t$, and $\Lambda(t') = \{\lambda'\}$ and $\Lambda(t) = \text{co}\{\lambda, \mu\}$ is the convex hull of λ' and μ. Since $\lambda' \succeq \lambda$, correspondence Λ obeys stochastic dominance in the sense given by (F1). By applying Proposition 3, we know that $\max \left\{ f(x, t') : x \in X \right\} \geq \max \left\{ f(x, t) : x \in X \right\}$, where f is as in (6). However, this does not guarantee that the optimal action is increasing, even when g is supermodular in (x, s). Indeed, let $g : X \times S \to \mathbb{R}$ be such that $g(0, s_1) = g(0, s_2) = 5$, $g(0, s_3) = 21$, $g(1, s_1) = 0$, $g(1, s_2) = 8$, and $g(1, s_3) = 24$, which is increasing in s and supermodular in (x, s). Since $\int_S g(0, s) d\lambda'(s) > \int_S g(1, s) d\lambda'(s)$, we have $\{0\} = \text{argmax} \left\{ f(x, t') : x \in X \right\}$.

Furthermore, since g is supermodular, we obtain $\int_S g(0,s)d\lambda(s) > \int_S g(1,s)d\lambda(s)$; however, this does not mean that the agent chooses action 0 at $\Lambda(t)$; In fact, since

$$\int_S g(0,s)d\lambda(s) > \int_S g(1,s)d\mu(s) > \int_S g(1,s)d\lambda(s) > \int_S g(0,s)d\mu(s),$$

it must be that $\{1\} = \operatorname{argmax}\left\{f(x,t) : x \in X\right\}$.

In order to guarantee that the agent finds it optimal to choose a higher action as beliefs shift, we need to formulate a condition for comparing sets of distributions that is more stringent than (F1). The following proposition provides such a necessary and sufficient condition and is the main result of this subsection.

Proposition 4. *Suppose that correspondence $\Lambda : T \to \triangle_S$ has compact and convex values. The following statements are equivalent.*

(i) *The correspondence Λ satisfies the following property:*

(F2) *for any $t' \geq t$, $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$, there is $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that*

$$\lambda' \succeq \mu, \quad \mu' \succeq \lambda, \quad \textit{and} \quad \tfrac{1}{2}\lambda + \tfrac{1}{2}\lambda' = \tfrac{1}{2}\mu + \tfrac{1}{2}\mu'.$$

(ii) *For any supermodular function $g : X \times S \to \mathbb{R}$, correspondence $\Gamma : X \times T \to \mathbb{R}^\ell$,*

$$\Gamma(x,t) \;:=\; \left\{a \in \mathbb{R}^\ell : a_i = -\delta_i(x)\lambda(s_i), \text{ for all } i = 1,\ldots,\ell, \text{ where } \lambda \in \Lambda(t)\right\}$$

is lower supermodular, where $\delta_i(x) = \left[g(x,s_{i+1}) - g(x,s_i)\right]$, for all $i = 1,\ldots,\ell$.

(iii) *The function $f : X \times T \to \mathbb{R}$, given by (6), is supermodular in (x,t), for any supermodular function $g : X \times S \to \mathbb{R}$.*

Remark 1. As shown in the proof, the claim (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) does not require Λ to have compact or convex values. These assumptions are used to prove that (iii) $\Rightarrow$ (i).

Remark 2. We show in the Appendix that Proposition 4 remains true if S is an interval of $\mathbb{R}$ and function $g(x,\cdot)$ is Riemann-Stieltjes integrable with respect to each $\lambda \in \Lambda(t)$, for all $x \in X$ and $t \in T$. This holds if any of the following conditions is satisfied: (a) function $g(x,s)$ is continuous in $s \in S$; (b) $g(x,s)$ is bounded on S and has only finitely many discontinuities in s, and all distributions in $\Lambda(t)$ are atomless; or (c) $g(x,s)$ is bounded on S and monotone, and all distributions in $\Lambda(t)$ are atomless.

Remark 3. The claim in this proposition remains true even if we leave part (i) unchanged; replace "lower supermodularity" in part (ii) with "upper supermodularity"; and replace the "min" operator in part (iii) with "max". In other words, (F2) is necessary and sufficient to guarantee that the function $F : X \times T \to \mathbb{R}$ given by

$$F(x,t) \;:=\; \max\left\{\int_S g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\right\} \tag{7}$$

is supermodular, for any supermodular function g. We prove this claim in the Appendix. The α-maxmin model by Ghirardato, Maccheroni, and Marinacci (2004) allows for both ambiguity averse and ambiguity loving behavior, with the agent's utility of the form

$$\alpha\;\min\left\{\int_S g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\right\} \;+\; (1-\alpha)\;\max\left\{\int_S g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\right\},$$

for some $\alpha \in [0,1]$. This function is supermodular in (x,t) if Λ satisfies (F2), since both elements of the sum are supermodular in (x,t). By the MCS theorem, the optimal action will also increase with t in the strong set order.

Proof of Proposition 4. The proof that (iii) implies (i) is in the Appendix. To show that (i) implies (ii), take any $x' \geq x$, $t' \geq t$ and $a \in \Gamma(x,t)$, $a' \in \Gamma(x',t')$. By definition of the correspondence Γ, there are distributions $\lambda \in \Lambda(t)$, $\lambda' \in \Lambda(t')$ such that $a_i = -\delta_i(x)\lambda(s_i)$ and $a'_i = -\delta_i(x')\lambda'(s_i)$, for all $i = 1,\ldots,\ell$. By (F2), there is $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that $\lambda'(s_i) \leq \mu(s_i)$, $\mu'(s_i) \leq \lambda(s_i)$, and $\lambda(s_i) + \lambda'(s_i) = \mu(s_i) + \mu'(s_i)$, for $i = 1,\ldots,\ell$. Since $g : X \times S \to \mathbb{R}$ is supermodular if and only if $\delta_i(x)$ is increasing (for each i),

$$\delta_i(x')\big[\mu(s_i) - \lambda'(s_i)\big] \;\geq\; \delta_i(x)\big[\mu(s_i) - \lambda'(s_i)\big] \;=\; \delta_i(x)\big[\lambda(s_i) - \mu'(s_i)\big],$$

for $i = 1,\ldots,\ell$. Construct vectors b, b', where $b_i = -\delta_i(x')\mu(s_i)$ and $b'_i = -\delta_i(x)\mu'(s_i)$, for all i. Clearly, $b \in \Gamma(x',t)$, $b' \in \Gamma(x,t')$, and $a + a' \geq b + b'$.

To show (ii) $\Rightarrow$ (iii), note that for any function $g : X \times S \to \mathbb{R}$ and distribution λ,

$$\begin{aligned}
\int_S g(x,s)d\lambda(s) &= g(x,s_1)\lambda(s_1) + \sum_{i=1}^{\ell} g(x,s_{i+1})\big[\lambda(s_{i+1}) - \lambda(s_i)\big] \\
&= g(x,s_{\ell+1})\lambda(s_{\ell+1}) + \sum_{i=1}^{\ell}\big[g(x,s_i) - g(x,s_{i+1})\big]\lambda(s_i) \qquad (8) \\
&= g(x,s_{\ell+1}) + \sum_{i=1}^{\ell}\big[-\delta_i(x)\lambda(s_i)\big],
\end{aligned}$$

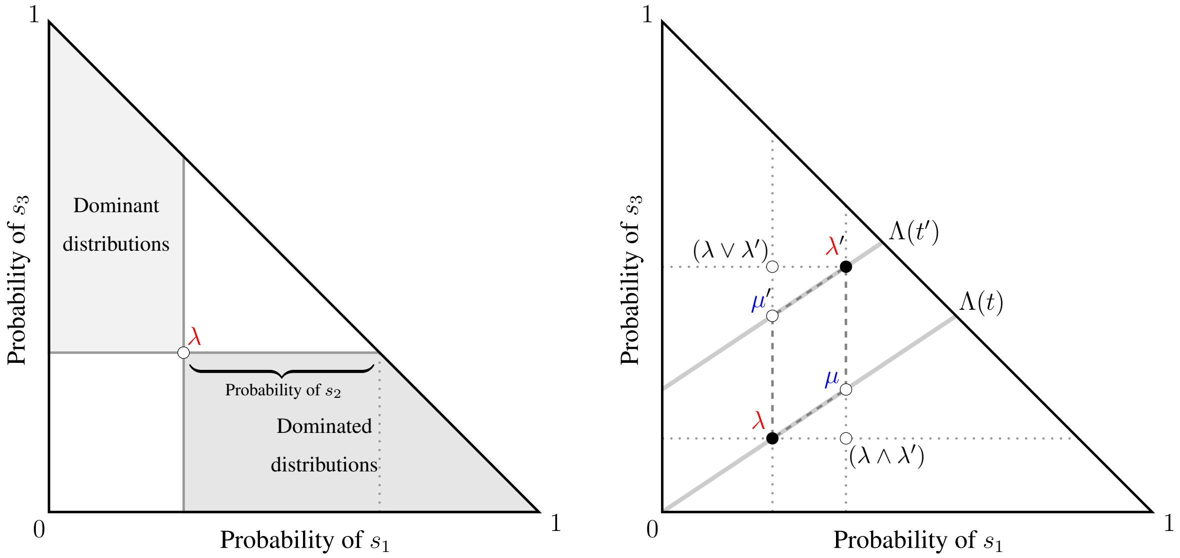

Figure 2: Probability measures represented in the Machina-Marshak triangle. On the right, the thick straight lines represent values $\Lambda(t)$ and $\Lambda(t')$ from Example 7.

since $\lambda(s_{\ell+1}) = 1$. Therefore, function f can be reformulated as

$$f(x,t) \;=\; g(x, s_{\ell+1}) + \min \Big\{ \mathbf{1} \cdot a : a \in \Gamma(x,t) \Big\},$$

where $\mathbf{1}$ is the unit vector and Γ is defined as in (ii). Since Γ is lower supermodular, the Main Theorem guarantees that f is a supermodular function. **QED**

We now have two intuitive set extensions of the notion of first order stochastic dominance: (F1) ensures monotone utility comparisons and (F2) monotone comparative statics. Clearly, (F2) implies (F1), but the converse is not true as shown in Example 5. Indeed, if we take the $\Lambda(t') = \{\lambda'\}$ and $\Lambda(t) = \text{co}\{\lambda, \mu\}$, then (F2) fails since (for example) there is no distribution in $\Lambda(t')$ that dominates $\mu \in \Lambda(t)$.

When does Λ satisfy (F2)? An obvious but restrictive example is when every distribution in $\Lambda(t')$ dominates every distribution in $\Lambda(t)$ if $t' > t$. The following examples give more general conditions under which (F2) holds.

Example 6 (Strong set order)**.** Suppose that the correspondence Λ is increasing in the *strong set order* induced by the first order stochastic dominance $\succeq$, i.e., for any $t' \geq t$, $\lambda \in \Lambda(t)$, and $\lambda' \in \Lambda(t')$, we have $\lambda \wedge \lambda' \in \Lambda(t)$ and $\lambda \vee \lambda' \in \Lambda(t')$. Since $\lambda' \succeq \lambda \wedge \lambda'$, $\lambda \vee \lambda' \succeq \lambda$, and $(\lambda \wedge \lambda') + (\lambda \vee \lambda') = \lambda + \lambda'$, the condition (F2) is satisfied. For example, let $\nu(\cdot, t)$ and $\mu(\cdot, t)$ be distributions in $\triangle_S$ that are increasing in t with respect to first order stochastic dominance and satisfy $\nu(\cdot,t) \succeq \mu(\cdot,t)$ for all t. Then the correspondence $\Lambda(t) = \{\lambda \in \triangle_S : \nu(\cdot,t) \succeq \lambda \succeq \mu(\cdot,t)\}$ that maps t to all distributions lying between $\nu(\cdot,t)$ and $\mu(\cdot,t)$ increases with t in the strong set order.

Example 7 (Increasing mean)**.** Take an increasing function $h : S \to \mathbb{R}$ and suppose that, for each $t \in T \subset \mathbb{R}$, the set $\Lambda(t)$ consists of all distributions over S for which the expected value of h is equal to t. Formally, let

$$\Lambda(t) = \left\{ \lambda \in \triangle_S : \int_S h(s) d\lambda(s) = t \right\}.$$

We show in the Appendix that $\Lambda : T \to \triangle_S$ satisfies (F2), even though it is clear that Λ is *not* increasing in the strong set order (see Figure 2 on the right).

In certain applications, it is natural for $g(x,s)$ to be increasing in s for all $x \in X$. In this case, one can assume, without loss of generality, that the belief correspondence Λ is upward comprehensive, i.e., if $\lambda \in \Lambda(t)$ and $\lambda' \succeq \lambda$ then $\lambda' \in \Lambda(t)$.[17] The following result (which we prove in the Appendix) states that when Λ is upward comprehensive, property (F2) remains a necessary and sufficient condition even if we only require the supermodularity of f for those functions g that are supermodular *and* increasing in s.[18]

Proposition 5. *Suppose that correspondence $\Lambda : T \to \triangle_S$ has compact, convex, and upper comprehensive values. Then the following statements are equivalent.*

(i) *The correspondence Λ satisfies property (F2).*

(ii) *The function $f(x,t)$, defined in (6), is supermodular in (x,t) for all functions g that are supermodular in (x,s) and increasing in s.*

We conclude this subsection with three economic applications.

Example 8 (Optimal savings)**.** Returning to the example in the Introduction, an agent decides on her savings $x \in X$ in period 1, given uncertainty on period 2 income $s \in S$. In that case, g is submodular in (x,s) and given by (1). Thus, if correspondence Λ increases in the sense of (F2), the agent will find it optimal to *reduce* savings.

[17] Given a correspondence Λ, let $\bar{\Lambda}(t) = \{\lambda \in \triangle_S : \lambda \succeq \lambda', \text{ for } \lambda' \in \Lambda(t)\}$. It is clear that $\bar{\Lambda}$ is upward comprehensive and that $\min\{\int_S g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\} = \min\{\int_S g(x,s)d\lambda(s) : \lambda \in \bar{\Lambda}(t)\}$, for all x.

[18] Notice that the comparative statics problem is dramatically simplified if g is increasing in s *and* $\Lambda(t)$ contains its infimum, i.e., a distribution $\underline{\lambda}(t)$ that is dominated by every other distribution in $\Lambda(t)$. Then $f(x,t) = \int_S g(x,s)d\underline{\lambda}(s)$, for all x, and we know that f is supermodular in (x,t) if $\underline{\lambda}(\cdot,t)$ increases with t. (This is consistent with Proposition 5 because $\Lambda(t) = \{\lambda \in \triangle_S : \lambda \succeq \underline{\lambda}(t)\}$ satisfies (F2).) However, there are natural examples of Λ obeying (F2) without $\Lambda(t)$ containing its infimum; see Example 7.

Example 9 (Portfolio problem)**.** An investor divides her wealth $m > 0$ between a *safe asset*, that pays out $r > 0$ for sure, and a *risky asset* with an uncertain gross payout of s in $S \subseteq \mathbb{R}_+$. The investor's beliefs over the risky return is captured by the correspondence $\Lambda : T \to \triangle_S$, where $\triangle_S$ is the space of probability distributions over S.

The investor chooses to invest $x \in X \subseteq \mathbb{R}$ in the risky asset, with the rest of her wealth invested in the safe security. We allow the investor to go short on either asset but require her to be solvent, i.e., it must be that $xs + (m - x)r \geq 0$, for all $s \in S$ and $x \in X$. Assuming that her Bernoulli index is $u : \mathbb{R}_+ \to \mathbb{R}$ and the investor is ambiguity averse, the investor's utility at $x \in X$ is given by

$$f(x, t) \ := \ \min \left\{ \int_S u\big(xs + (m - x)r\big) d\lambda(s) : \lambda \in \Lambda(t) \right\}. \qquad (9)$$

To capture the idea that a higher t represents greater optimism, we assume that correspondence Λ increases in t according to (F2). In particular, this implies that the function f is supermodular if $g(x, s) := u\big(xs+(m-x)r\big)$ is supermodular. Assuming that u is strictly increasing, concave, and twice continuously differentiable, it is straightforward to check that g is supermodular if the coefficient of relative risk aversion of u is less than 1.[19] Therefore, with this condition on u, we can apply the MCS theorem to guarantee that the investor's holding in the risky asset increases with t. This conclusion holds even if the investor's preference has the α-maxmin form.[20]

The next example has a different flavor from Example 9: it has both x and t as choice variables and exploits the fact that supermodularity is preserved by the sum.

Example 10. A firm operating in uncertain market conditions must decide on how much to produce and how much to spend on promoting its product via advertising. In period 1, the marginal cost of production is $c > 0$ and the marginal cost of advertising is $a > 0$. If the firm chooses t units of advertising, its belief on the demand for its output s is given by a set of distributions $\Lambda(t) \subseteq \triangle_S$; higher advertising leads to greater demand in the sense that Λ satisfies (F2). We assume that the price of the good is fixed at 1.

[19]Note that, since x can take negative values, function g does not increase in s.

[20] We are not the first to discuss comparative statics of the portfolio choice model under ambiguity. For example, Gollier (2011) examines how the demand for the risky asset changes with the level of ambiguity aversion, in the context of the smooth ambiguity model. Cherbonnier and Gollier (2015) study both the smooth ambiguity model and the α-maxmin model; the authors provide conditions under which the demand for the risky asset increases with respect to initial wealth.

In period 2, the firm's actual demand s is realized and the firm has to meet this demand even if it exceeds its period 1 output; the profit in period 2 is

$$\pi(x,s) \;:=\; s - \kappa\big(\max\{s-x, 0\}\big).$$

Function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ should be interpreted as the cost of producing the additional units to meet demand in period 2. At the same time, goods for which there is no demand can be freely disposed. Also, notice that $\pi(x,s)$ need not be increasing in s.

The firm chooses $x \geq 0$ and $t \geq 0$ in period 1 to maximise

$$\Pi(x,t,c,a) \;:=\; \min\left\{\int_S \pi(x,s)d\lambda(s) : \lambda \in \Lambda(t)\right\} - cx - at.$$

It is straightforward to check that the function π is supermodular if κ is increasing, convex, and $\kappa(0) = 0$.[21] Given this, Proposition 4 guarantees that

$$f(x,t) \;=\; \min\left\{\int_S \pi(x,s)d\lambda(s) : \lambda \in \Lambda(t)\right\}$$

is a supermodular function of (x,t) and therefore Π is supermodular in (x,t). Furthermore, Π has increasing differences in $\big((x,t),(-c,-a)\big)$. Applying the MCS theorem, we conclude that more advertising and higher output will ensue from either a fall in the cost of advertising a or a fall in the cost of period 1 production c.

5.2 Variational and multiplier preferences

Proposition 4 can be extended to cover a broader class of preferences. Maccheroni, Marinacci, and Rustichini (2006) introduce and axiomatize a generalization of the maxmin model, called *variational preferences*. In this model, the utility of some action x is $f(x) = \min\left\{\int_S g(x,s)d\lambda(s) + c(\lambda) : \lambda \in \triangle_S\right\}$. Loosely speaking, the agent's utility from action x is obtained by minimizing her expected utility over the set of all probability distributions; unlike the maxmin model where the agent is restricted to a subset of $\triangle_S$, any distribution in $\triangle_S$ could be 'picked' in the variational preferences model, though each distribution λ is associated with a different cost $c(\lambda)$.[22] Below, we parameterize the

[21] Take any $x' \geq x$ and consider three cases. If (i) $s \leq x$, then $\delta(s) := \big[\pi(x',s) - \pi(x,s)\big] = 0$; whenever (ii) $x < s \leq x'$, then $\delta(s) = \kappa(s-x)$; and finally (iii) $s > x'$ implies $\delta(s) = \kappa(s-x) - \kappa(s-x')$. In either case, under the assumptions imposed on κ, the function δ is increasing in s.

[22] For a discussion see Maccheroni, Marinacci, and Rustichini (2006) or Epstein and Schneider (2010).

cost function c by $t \in T \subseteq \mathbb{R}$ and identify conditions under which the agent's utility is supermodular in (x, t).

Proposition 6. *Let $c : \triangle_S \times T \to \mathbb{R}_+$ be a continuous and convex function on $\triangle_S$, for all $t \in T$. The following statements are equivalent.*

(i) *The function c satisfies the following property:*

(C) *for any $t' \geq t$ in T and λ, λ' in $\triangle_S$ there is μ, μ' in $\triangle_S$ such that*

$$\lambda' \succeq \mu,\ \mu' \succeq \lambda,\ \tfrac{1}{2}\lambda + \tfrac{1}{2}\lambda' = \tfrac{1}{2}\mu + \tfrac{1}{2}\mu',\ \text{and } c(\lambda, t) + c(\lambda', t') \geq c(\mu, t) + c(\mu', t').$$

(ii) *For any supermodular function $g : X \times S \to \mathbb{R}$, correspondence $\Gamma : X \times T \to \mathbb{R}^{\ell+1}$,*

$$\Gamma(x, t) := \left\{ a \in \mathbb{R}^{\ell+1} : a_i = -\delta_i(x)\lambda(s_i),\ \text{if } i = 1, \dots, \ell,\ \text{and } a_{\ell+1} = c(\lambda, t),\ \text{for } \lambda \in \triangle_S \right\}$$

is lower supermodular, where $\delta_i(x) = \big[g(x, s_{i+1}) - g(x, s_i)\big]$, for $i = 1, \dots, \ell$.

(iii) *Function $f : X \times T \to \mathbb{R}$, where*

$$f(x, t) \ := \ \min \left\{ \int_S g(x, s) d\lambda(s) + c(\lambda, t) : \lambda \in \triangle_S \right\}, \tag{10}$$

is supermodular for any supermodular function $g : X \times S \to \mathbb{R}$.

The proof is found in the Appendix. Implication (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) does not require the cost function c to be convex or continuous. We employ the additional assumption to prove that (iii) $\Rightarrow$ (i). Condition (C) in the proposition can be thought of as generalization of condition (F2) imposed on $\Lambda : T \to \triangle_S$ in Proposition 6. Indeed, given Λ, we define

$$c(\lambda, t) \ = \ \begin{cases} 0 & \text{if } \lambda \in \Lambda(t); \\ \infty & \text{otherwise.} \end{cases}$$

Then c obeys (C) if and only if Λ obeys (F2), while (10) reduces to the maxmin form (6) in this case. Below are two more examples of cost functions that satisfy property (C).

Example 11 (Submodular cost)**.** Suppose that function $c : \triangle_S \times T \to \mathbb{R}_+$ is submodular, i.e., for all λ, $\lambda' \in \triangle_S$ and t, $t' \in T$,

$$c(\lambda, t) + c(\lambda', t') \ \geq \ c(\lambda \vee \lambda', t \vee t') + c(\lambda \wedge \lambda', t \wedge t').$$

Then c obeys condition (C), as we can always choose $\mu = \lambda \wedge \lambda'$ and $\mu' = \lambda \vee \lambda'$.

Example 12. Suppose that $\tilde{c} : \mathbb{R} \times T \to \mathbb{R}$ is a submodular function and the cost function $c : \triangle_S \times T \to \mathbb{R}$ is evaluated by $c(\lambda, t) := \tilde{c}\left(\int_S h(s) d\lambda(s), t\right)$ for some increasing function $h : S \to \mathbb{R}$. In other words, the cost function depends only on the mean of the random variable h with respect to the distribution λ, and the parameter t. We claim that c satisfies condition (C). Indeed, take any λ, λ' in $\triangle_S$ and denote the mean of function h corresponding to each distribution by m, m' respectively. Suppose that $m' \geq m$; then there are distributions μ, μ' with means m, m' respectively, such that $\lambda' \succeq \mu$, $\mu' \succeq \lambda$, and $(1/2)\lambda + (1/2)\lambda' = (1/2)\mu + (1/2)\mu'$.[23] Since $c(\lambda, t) = c(\mu, t)$ and $c(\lambda', t') = c(\mu', t')$, we obtain $c(\lambda, t) + c(\lambda', t') = c(\mu, t) + c(\mu', t')$. If $m' < m$, choose $\mu = \lambda'$ and $\mu' = \lambda$. By the submodularity of $\tilde{c}$, we obtain (as required)

$$c(\lambda, t) + c(\lambda', t') \;=\; \tilde{c}(m, t) + \tilde{c}(m', t') \;\geq\; \tilde{c}(m', t) + \tilde{c}(m, t') \;=\; c(\mu, t) + c(\mu', t').$$

An important sub-class of variational preferences are *multiplier preferences*, which were used in Sargent and Hansen (2001) and axiomatized by Strzalecki (2011a). In this case, the cost function is $c(\lambda, t) = \theta R\big(\lambda \| \lambda^*(\cdot, t)\big)$, for $\theta \geq 0$ and $\lambda^*(\cdot, t) \in \triangle_S$, where

$$R\big(\lambda \| \lambda^*(\cdot, t)\big) \;:=\; \int_S \ln\left(\frac{d\lambda(s)}{d\lambda^*(s, t)}\right) d\lambda(s)$$

is the *relative entropy*.[24] Note that $d\lambda(s)$, $d\lambda^*(s, t)$ denote the probability of state s in the distribution λ, $\lambda^*(\cdot, t)$, respectively. This representation can be interpreted in the following manner. The decision maker has a belief over the states of the world given by a *reference* or *benchmark* distribution $\lambda^*(\cdot, t)$, but she is not completely confident that she is exactly correct. To accommodate this concern, the decision maker takes all distributions in $\triangle_S$ into account when evaluating her utility from a given action, though distributions further away from $\lambda^*(\cdot, t)$ cost more and are thus less likely to be the distribution that solves the minimization problem in (10).

The multiplier preferences model has a cost function that is particularly well-behaved.

Proposition 7. *The cost function $c : \triangle_S \times T \to \mathbb{R}$, given by $c(\lambda, t) := \theta R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular on $\triangle_S$, for all $t \in T$ and positive θ. Furthermore, if $\lambda^*(\cdot, t)$ is increasing in t with respect to the monotone likelihood ratio,*[25] *then c is submodular in (λ, t).*

[23] For a proof of this claim, see the proof of Example 7 in the Appendix.

[24] See Strzalecki (2011b) for a detailed discussion on the relation between variational preferences, multiplier preference, and subjective expected utility.

[25] This requires that, for any $t' \geq t$, the ratio $d\lambda^*(s, t')/d\lambda^*(s, t)$ be increasing with s. This property implies $\lambda^*(\cdot, t') \succeq \lambda^*(\cdot, t)$.

We prove this result in the Appendix. Note that when $c(\lambda, t)$ is submodular in (λ, t) then it obeys condition (C) (see Example 11). By applying Proposition 6, we conclude that $f(x, t)$ is supermodular in (x, t) if $g(x, s)$ is supermodular in (x, s) and $\lambda^*(\cdot, t)$ is increasing in t with respect to the monotone likelihood ratio order. This result captures the idea that as the agent revises her benchmark belief towards higher states, the cost function changes in a way that raises the marginal utility to her of taking higher actions.

In Examples 8, 9 and 10, we gave economic applications of Proposition 4, assuming that the agent has maxmin utility. It is clear that, by appealing to Proposition 6, the conclusions in those examples will continue to hold, mutatis mutandi, if the agent has variational or, more specifically, multiplier preferences.

6 Dynamic programming under ambiguity aversion

In an influential paper, Hopenhayn and Prescott (1992) used the tools of monotone comparative statics to analyze stationary dynamic optimization problems. In this section, we show how those results could be extended to the case where the agent has a multi-prior belief, by applying the results from the previous section.

We consider an agent who faces a stochastic control problem where X and S are the sets of endogenous and exogenous state variables, respectively. To keep the exposition simple, we shall assume that X is a sublattice of a Euclidean space and S is a subset of another Euclidean space. The evolution of s over time follows a Markov process with the transition function λ. The agent's problem can be formulated in the following way (see Stokey, Lucas, and Prescott, 1989). At each period τ, given the current state $(x_\tau, s_\tau) \in X \times S$, the agent chooses the endogenous variable $x_{\tau+1}$ for the following period. We assume that $x_{\tau+1}$ is chosen from a non-empty feasible set which may depend on the current state and which we denote by $B(x_\tau, s_\tau) \subseteq X$. The single-period return is given by the function $F : X \times S \times X \to \mathbb{R}$; $F(x, s, y)$ is the payoff when (x, s) is the state variable in period τ and y is the endogenous state variable in period $\tau + 1$ chosen in period τ. The agent discounts these payoffs by a constant factor $\beta \in (0, 1)$.

The agent's objective is to maximize her expected discounted payoffs over an infinite horizon, given the initial condition (x, s). We denote the value of this optimization problem by $v^*(x,s)$. Under standard assumptions — in particular, the continuity and boundedness of F and the continuity of B — this problem admits a recursive representation, where $v = v^*$ is the unique solution to the Bellman equation

$$v(x,s) \;=\; \max\left\{F(x,s,y) + \beta \int_S v(y,\tilde{s})d\lambda(\tilde{s},s) : y \in B(x,s)\right\},$$

where $\lambda(\cdot,s)$ is a cumulative probability distribution over states of the world in the following period, conditional on the current state s.[26] The function v^* is bounded and continuous. Moreover, the operator $\mathscr{T} : \mathcal{B} \to \mathcal{B}$ defined by

$$\big(\mathscr{T}v\big)(x,s) \;=\; \max\left\{u(x,s,y) + \beta \int_S v(y,\tilde{s})d\lambda(\tilde{s},s) : y \in B(x,s)\right\},$$

that maps the space $\mathcal{B}$ of bounded and continuous real-valued functions over $X \times S$ into itself has the following property: beginning at *any* $v \in \mathcal{B}$, the function $(\mathscr{T}^n v)$ converges uniformly to v^* as n tends to infinity.[27] Furthermore, the set

$$\Phi(x,s) \;:=\; \arg\max\left\{F(x,s,y) + \beta \int_S v^*(y,\tilde{s})d\lambda(\tilde{s},s) : y \in B(x,s)\right\}$$

is non-empty and compact, for all $(x,s) \in X \times S$, and the correspondence $\Phi : X \times S \to X$ is upper hemi-continuous. We refer to any optimal control problem in which v^* and Φ have the properties listed in this paragraph as a *well-behaved* problem.

Given a well-behaved problem, Hopenhayn and Prescott (1992) (henceforth HP) apply Theorem 4.3 in Topkis (1978) to show that the value $v^*(x,s)$ *is supermodular in x and has increasing differences in (x,s)* under the following assumptions: (i) $F(x,s,y)$ is supermodular in (x,y) and has increasing differences in $\big((x,y),s\big)$; (ii) the graph of B is a sublattice of $X \times S \times X$; (iii) $\lambda(\cdot,s)$ is increasing in s with respect to the first order stochastic dominance. The properties of v^* in turn guarantee that the function

$$f(x,s,y) \;:=\; F(x,s,y) + \beta \int_S v^*(y,\tilde{s})d\lambda(\tilde{s},s)$$

is supermodular in y and has increasing differences in $\big(y,(x,s)\big)$. By the MCS theorem, $\Phi(x,s)$ is a compact sublattice of X and is increasing in (x,s) in the strong set order.[28]

[26] See Theorem 9.6 in Stokey, Lucas, and Prescott (1989) for details.

[27] $\mathscr{T}^n$ is the nth orbit of the operator $\mathscr{T}$, i.e., $(\mathscr{T}^{n+1}v) = \mathscr{T}\big(\mathscr{T}^n v\big)$.

[28] Condition (ii) on B guarantees that $B(x,s)$ is sublattice of X and that it increases with (x,s) in the strong set order. Given the properties on f, we know that $\Phi(x,s)$ is a sublattice and that it increases with (x,s) in the strong set order; this follows from a more general version of the MCS theorem (than the one stated in Section 2) that allows for increasing constraint sets. See Topkis (1978).

This in turn guarantees that the greatest optimal selection

$$\phi(x,s) \;:=\; \Big\{ y \in \Phi(x,s) : y \geq_X z, \text{ for all } z \in \Phi(x,s) \Big\},^{29}$$

exists and is an increasing and Borel measurable of (x, s). Lastly, the policy function ϕ induces a Markov process on $X \times S$, where, for measurable sets $Y \subseteq X$ and $T \subseteq S$, the probability of $Y \times T$ conditional on (x, s) is the probability of T conditional on s if $\phi(x, s) \in Y$, and it is zero otherwise. HP make use of the monotonicity of ϕ to guarantee that this Markov process has a stationary distribution.[30]

We now consider a stochastic control problem identical to the one we just described, except that we allow the agent to be ambiguity averse. Since at each period τ the exogenous variable is drawn from the set S, the set of all possible realizations of the exogenous variable over time is given by S^∞. An expected utility maximizer behaves as though she is guided by a distribution over S^∞; to obtain the utility of a given plan of action, the agent evaluates the discounted utility on every possible path, i.e., over every element in S^∞ and takes the average across paths, weighing each path with its probability. When the agent has a maxmin preference, her behavior can be modelled by a *set* of distributions $\mathcal{M}$ over S^∞. The utility of a plan is then given by the minimum of the expected discounted utility for every distribution in $\mathcal{M}$.

In contrast to expected discounted utility, it is known that the agent's utility in the maxmin model will not generally have a recursive representation. However, there is a condition on $\mathcal{M}$ called *rectangularity* which is sufficient (and effectively necessary) for this to hold (see Epstein and Schneider, 2003). Furthermore, it is known that a time-invariant version of rectangularity is also sufficient to guarantee that the agent's problem can be solved through the Bellman equation, in a way analogous to that for expected discounted utility (see Iyengar, 2005). This condition says that the agent's belief over the possible value of the exogenous variable in the following period, after observing s in the current period, is given by a *set* of distribution functions $\Lambda(s)$; this set depends on the current realization of the exogenous variable and is time-invariant. The set $\mathcal{M}$,

[29] Function is well-defined because Φ is compact-valued and a sublattice.

[30] The focus in this section is on primitive conditions guaranteeing the monotonicity of the policy function. Readers who are interested in how the distribution over (x, s) evolves over time (under monotonicity or weaker assumptions) should consult Huggett (2003). HP and Stachurski and Kamihigashi (2014) also discuss uniqueness and other issues relating to the stationary distribution.

given an initial value s_0, is then obtained by concatenating the transition probabilities. Therefore, the probability associated with a path $(s_0, s_1, s_2, s_3, \ldots)$ is $\prod_{i=1}^{\infty} p_i$, where p_1 is the probability of s_1 for some distribution in $\Lambda(s_0)$, p_2 is the probability of s_2 for some distribution in $\Lambda(s_1)$, etc.

With this assumption on $\mathcal{M}$ in place, and some other standard conditions, one could guarantee that the value $v^*(x, s)$ of the control problem with the initial state (x, s), is the unique solution to the Bellman equation

$$v(x, s) \;=\; \max \Big\{ F(x, s, y) + \beta(Av)(y, s) : y \in B(x, s) \Big\}$$

where $(Av)(y, s) = \min \big\{ \int_S v(y, \tilde{s}) d\lambda(\tilde{s}) : \lambda \in \Lambda(s) \big\}$ (see Iyengar, 2005). Furthermore, the problem is *well-behaved* in the sense defined at the beginning of this section, i.e., the operator $\mathscr{T}$ given by $(\mathscr{T}v)(x, s) := \max \Big\{ u(x, s, y) + \beta(Av)(y, s) : y \in B(x, s) \Big\}$ converges uniformly to v^* and the correspondence Φ is upper hemi-continuous on $X \times S$.

With this basic set-up, we are almost in a position to recover a monotone result of the HP type: all that is needed is a condition guaranteeing that $(Av)(y, s)$ is a supermodular function of (y, s), whenever v is supermodular. When X and S are one-dimensional, Proposition 4 tells us that this holds if the beliefs correspondence Λ satisfies (F2). The proof of the next proposition is supplied in the Appendix.

Proposition 8. *Consider a well-behaved optimal control problem where X, $S \subset \mathbb{R}$, with X compact and S finite. Suppose that $F(x, s, y)$ is supermodular in (x, s, y), $\Lambda : S \to \triangle_S$ satisfies (F2), and the graph of $B : X \times S \to X$ is a sublattice; then the value function $v^*(x, s)$ is supermodular in (x, s) and the correspondence $\Phi : X \times S \to \mathbb{R}$, where*

$$\Phi(x, s) \;:=\; \arg\max \Big\{ F(x, s, y) + \beta(Av^*)(y, s) : y \in B(x, s) \Big\},$$

is sublattice-valued and increasing in the strong set order. Moreover, the greatest selection $\phi : X \times S \to \mathbb{R}$ of Φ is well-defined, increasing, and Borel measurable.

Below we discuss an application of this result.

Example 13. Consider the following dynamic optimization problem of a firm. In each period, the firm collects revenue $\pi(x, s)$, where $s \in S$ denotes the realized exogenous state of the world and $x \in \mathbb{R}_+$ is the level of capital stock currently available to the firm. Once

s is revealed to the firm and the revenue collected, the firm may invest $a \in [0, K]$ at a cost $c(a)$, K being a finite positive number. With this investment, capital stock in the next period is $y = \delta x + a$, where $\delta \in [0, 1]$ denotes the fraction of non-depreciated capital. Therefore, the dividend in each period is

$$F(x, s, y) \; := \; \pi(x, s) - c(y - \delta x),$$

where the firm chooses y from the interval $B(x, s) = [\delta x, \delta x + K]$. We know from HP that if the firm is an expected utility maximizer and the optimal control problem is well-behaved, then the firm has a policy function that is increasing in (x, s) under the following additional conditions: the transition *function* $\Lambda : S \to \triangle_S$ is increasing with respect to first order stochastic dominance and F is supermodular; the latter property is guaranteed if π is supermodular in (x, s) (but not necessarily increasing in s) and c is concave. Proposition 8 goes further by saying that this conclusion remains true if the firm has a maxmin preference, so long as the transition *correspondence* Λ satisfies (F2).

Appendix

Proof of Proposition 1. We show part (i) and omit the proof of (ii), which is analogous. Towards contradiction, suppose that Γ is not upper supermodular. Hence, there is x, x' in X and $y \in \Gamma(x)$, $y' \in \Gamma(x')$ such that for any $z \in \mathrm{co}\Gamma(x \wedge x')$ and $z' \in \mathrm{co}\Gamma(x \vee x')$, we have $z + z' \not\geq y + y'$.

Let $U = \Big\{u \in Y : u \leq v, \text{ for some } v \in \mathrm{co}\Gamma(x \wedge x') + \mathrm{co}\Gamma(x \vee x')\Big\}$. This set is convex, downward comprehensive, and $(y + y') \notin U$. Moreover, $\mathrm{co}\Gamma(x \wedge x') + \mathrm{co}\Gamma(x \vee x')$ is a closed set by assumption, and therefore, so is U. By the strong separating hyperplane theorem, there is a non-zero, linear functional ϕ^* such that $\phi^*(y + y') > \phi^*(u)$, for all $u \in U$. As U is downward comprehensive, ϕ^* must be positive.

We claim that $f(x) = \max\big\{\phi^*(u) : u \in \Gamma(x)\big\}$ is *not* supermodular. Indeed,

$$\begin{aligned} f(x \wedge x') + f(x \vee x') \;&=\; \max\big\{\phi^*(u) : u \in \Gamma(x \wedge x')\big\} + \max\big\{\phi^*(u) : u \in \Gamma(x \vee x')\big\} \\ &=\; \max\big\{\phi^*(u) : u \in \Gamma(x \wedge x') + \Gamma(x \vee x')\big\} \;<\; \phi^*(y + y') \;=\; \phi^*(y) + \phi^*(y') \\ &\leq\; \max\big\{\phi^*(u) : u \in \Gamma(x)\big\} + \max\big\{\phi^*(u) : u \in \Gamma(x')\big\} \;=\; f(x) + f(x'), \end{aligned}$$

which contradicts the supermodularity of f. **QED**

Proof of the claim in Example 4. Take any $t' \geq t$ and $q' \geq q$. We need to show that for all $y \in \Gamma(t', q)$, $y' \in \Gamma(t, q')$ there is some $z \in \Gamma(t, q)$ and $z' \in \Gamma(t', q')$ such that $z + z' \leq y + y'$. By definition, $F(t', y) \geq q$ and $F(t, y') \geq q'$. Since F is increasing in t, $F(t', y') \geq q'$; if, in addition, $F(t, y) \geq q$, then we can choose $z = y$ and $z' = y'$ and $z + z' = y + y'$ obviously holds.

Alternatively, suppose that $F(t, y) < q$. Let $v = y' - (y \wedge y')$. By monotonicity of F, we have $F(t, y \wedge y') \leq F(t, y) < q < q' \leq F(t, y')$. Since $F(t, \cdot)$ is continuous, there is $\lambda \geq 0$ such that $F\big(t, (y \wedge y') + \lambda v\big) = q$. Then

$$\begin{aligned} q' - q \ &\leq\ F(t, y') - F\big(t, (y \wedge y') + \lambda v\big) \ \leq\ F(t', y') - F\big(t', (y \wedge y') + \lambda v\big) \\ &\leq\ F\big(t', (y \vee y') - \lambda v\big) - F(t', y) \ \leq\ F\big(t', (y \vee y') - \lambda v\big) - q, \end{aligned}$$

where the second inequality holds because F obeys increasing differences between t and the input vector and the third inequality holds because $F(t', \cdot)$ is supermodular and concave. (For a proof of the second claim see Proposition 2 in Quah, 2007.) Hence $q' \leq F\big(t', (y \vee y') - \lambda v\big)$; in other words, $[(y \vee y') - \lambda v] \in \Gamma(t', q')$. Letting $z = (y \wedge y') + \lambda v$ and $z' = (y \vee y') - \lambda v$, we obtain $z + z' = (y \wedge y') + (y \vee y') = y + y'$. **QED**

Proof that (ii) $\Rightarrow$ (i) in Proposition 3. We prove this by contradiction. If (F1) fails, there is some $t' \geq t$ and $\lambda' \in \Lambda(t')$ such that $\lambda' \not\geq \lambda$, for all $\lambda \in \Lambda(t)$. Let $V = \big\{y \in \mathbb{R}^\ell : y_i \geq \lambda'(s_i), \text{ for } i = 1, \ldots, \ell\big\}$. Since $V \cap \Lambda(t') = \emptyset$ and $\big(V - \Lambda(t')\big)$ is closed and convex, by the strong separating hyperplane theorem, $\min\Big\{\sum_{i=1}^\ell \hat{p}_i y_i : y \in V\Big\} > \max\Big\{\sum_{i=1}^\ell \hat{p}_i \lambda(s_i) : \lambda \in \Lambda(t')\Big\}$, for some $\hat{p} \in \mathbb{R}^\ell$. Given that V is upward comprehensive, $\hat{p} > 0$; furthermore, $\sum_{i=1}^\ell \hat{p}_i \lambda'(s_i) = \min\big\{\hat{p} \cdot y : y \in V\big\}$. Define $u : S \to \mathbb{R}$ by $u(s_1) = \hat{p}_1$ and $u(s_{i+1}) = u(s_i) + \hat{p}_{i+1}$, for $i = 1, \ldots, \ell$. Note that u is an increasing function. Since $\int_S u(s) d\mu(s) = u(s_{\ell+1}) - \sum_{i=1}^\ell \hat{p}_i \mu(s_i)$ for any $\mu \in \triangle_S$ (recall (8) in Section 5.1),

$$\begin{aligned} \min\left\{\int_S u(s) d\lambda(s) : \lambda \in \Lambda(t)\right\} \ &=\ u(s_{\ell+1}) - \max\left\{\sum_{i=1}^\ell \hat{p}_i \lambda(s_i) : \lambda \in \Lambda(t)\right\} \\ &>\ u(s_{\ell+1}) - \sum_{i=1}^\ell \hat{p}_i \lambda'(s_i) \ \geq\ u(s_{\ell+1}) - \max\left\{\sum_{i=1}^\ell \hat{p}_i \lambda(s_i) : \lambda \in \Lambda(t')\right\} \\ &=\ \min\left\{\int_S u(s) d\lambda(s) : \lambda \in \Lambda(t')\right\}. \end{aligned}$$

Thus (F1) is indeed necessary for monotone maxmin utility. **QED**

Proof that (iii) ⇒ (i) in Proposition 4. Suppose Λ violates (F2), so that for some $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$ there is no $\mu \in \Lambda(t)$, $\mu' \in \Lambda(t')$ such that $\mu' \succeq \lambda$, $\lambda' \succeq \mu$, and $(1/2)\lambda + (1/2)\lambda' = (1/2)\mu + (1/2)\mu'$. Let

$$D \;=\; \Big\{(d,d') \in \mathbb{R}^\ell \times \mathbb{R}^\ell \;:\; d_i = \big[\mu'(s_i) - \lambda(s_i)\big] \text{ and } d'_i = \big[\mu(s_i) - \lambda'(s_i)\big], \\ \text{for } \; i = 1,\dots,\ell, \text{ where } \mu \in \Lambda(t) \text{ and } \mu' \in \Lambda(t')\Big\}. \quad (11)$$

Then it is clear that $D \cap C = \emptyset$, where $C = \Big\{(c,c') \in \mathbb{R}^\ell \times \mathbb{R}^\ell \;: c = -c' \text{ and } c' \in \mathbb{R}^\ell_+\Big\}$. Since C is closed, convex and contains $(0,0)$, and D is compact and convex, by the strong separating hyperplane theorem, there is $(\hat{p},\hat{p}')$ in $\mathbb{R}^\ell \times \mathbb{R}^\ell$ such that

$$\hat{p}\cdot d + \hat{p}'\cdot d' \;<\; 0 \;\leq\; \hat{p}\cdot c + \hat{p}'\cdot c',$$

for all $(c,c') \in C$ and $(d,d') \in D$. Let ϵ_i denote the ℓ-dimensional vector with all entries equal to 0 apart from ith-entry i, which equals 1. Given that $(-\epsilon_i,\epsilon_i)$ belongs to C, for all $i = 1,\dots,\ell$, we obtain $\hat{p}' \geq \hat{p}$. Take any $x, x' \in X$ such that $x' > x$ and define function $g : X \times S \to \mathbb{R}$ as follows. Let $g(y,s_1) = 0$, for all $y \in X$, and

$$g(y,s_i) \;:=\; \begin{cases} \sum_{j=1}^{i-1}\hat{p}_i & \text{if } y < x'; \\ \sum_{j=1}^{i-1}\hat{p}'_i & \text{otherwise,} \end{cases} \quad (12)$$

for all $i = 2,\dots,(\ell+1)$. The function g is supermodular because $\hat{p}' \geq \hat{p}$. Moreover, for any $\mu \in \Lambda(t)$, $\mu' \in \Lambda(t')$, we have

$$\int_S g(x,s)d\lambda(s) - \int_S g(x,s)d\mu'(s) + \int_S g(x',s)d\lambda'(s) - \int_S g(x',s)d\mu(s) \\ = \sum_{i=1}^{\ell} p_i\big[\mu'(s_i) - \lambda(s_i)\big] + \sum_{i=1}^{\ell} p'_i\big[\mu(s_i) - \lambda'(s_i)\big] \;<\; 0$$

since $(\mu' - \lambda, \mu - \lambda') \in D$. This holds for *any* $\mu \in \Lambda(t)$, $\mu' \in \Lambda(t')$; as a consequence,

$$f(x,t) + f(x',t') \;\leq\; \int_S g(x,s)d\lambda(s) + \int_S g(x',s)d\lambda'(s) \\ <\; \min\Big\{\int_S g(x,s)d\nu(s) : \nu \in \Lambda(t')\Big\} + \min\Big\{\int_S g(x',s)d\nu(s) : \nu \in \Lambda(t)\Big\} \\ =\; f(x,t') + f(x',t).$$

So f is not supermodular, contradicting (iii). **QED**

Proof of Remark 2 following Proposition 4. Suppose that $S = [a, b]$. Let $\{s_i^n\}_{i=0}^n$ be a sequence with $n+1$ terms such that $a = s_0^n < s_1^n < \ldots < s_{n-1}^n < s_n^n = b$. Since at each (x, t), function $g(x, \cdot)$ is the Riemann-Stieltjes integrable with respect to $\lambda \in \Lambda(t)$, we can choose $\{s_i^n\}_{i=0}^n$ so that

$$\int_S g(x,s)d\lambda(s) \;=\; \lim_{n\to\infty} \sum_{i=0}^{n-1} g(x, s_{i+1})\big[\lambda(s_{i+1}) - \lambda(s_i)\big]$$

for all $\lambda \in \Lambda(t)$. This guarantees that $\lim_{n\to\infty} f_n(x,t) = f(x,t)$ for all (x,t), where

$$f_n(x,t) \;:=\; \min\left\{\sum_{i=0}^{n-1} g(x, s_{i+1}^n)\big[\lambda(s_{i+1}^n) - \lambda(s_i^n)\big] : \lambda \in \Lambda(t)\right\}.$$

We know, from the case where S is finite, that $f_n : X \times T \to \mathbb{R}$ is a supermodular function. Since supermodularity is preserved by pointwise convergence, f is supermodular. **QED**

Proof of Remark 3 following Proposition 4. The proof mimics the one for Proposition 4 so we only provide a sketch. To show that (i) implies the upper supermodularity of Γ, take any $x' \geq x$, $t \geq t'$ and $a \in \Gamma(x', t)$, $a' \in \Gamma(x, t')$; by the definition of Γ, there is $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$ such that $a_i = -\delta_i(x')\lambda(s_i)$ and $a'_i = -\delta_i(x)\lambda'(s_i)$, for all i. By (F2) there is $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that $\lambda'(s_i) \leq \mu(s_i)$, $\mu'(s_i) \leq \lambda(s_i)$, and $\lambda(s_i) + \lambda'(s_i) = \mu(s_i) + \mu'(s_i)$, for all i. Since g is supermodular, $\delta_i(x')\big[\lambda(s_i) - \mu'(s_i)\big] \geq \delta_i(x)\big[\lambda(s_i) - \mu'(s_i)\big] = \delta_i(x)\big[\mu(s_i) - \lambda'(s_i)\big]$, for each i. Construct vectors b, b' where $b_i = -\delta_i(x)\mu(s_i)$ and $b'_i = -\delta_i(x')\mu'(s_i)$, for all i. Clearly, $b \in \Gamma(x', t)$, $b' \in \Gamma(x, t')$, and $b + b' \geq a + a'$.

The proof that the upper supermodularity of Γ implies that F, given by (7), is supermodular is a straightforward application of the Main Theorem and we shall omit it. Lastly, we show that if (F2) is violated, then there is a supermodular function g for which F is *not* supermodular. As in the proof that (iii) implies (i) in Proposition 4, we first obtain $\hat{p}$ and $\hat{p}'$ such that $\hat{p} \leq \hat{p}'$ and $\hat{p} \cdot d + \hat{p}' \cdot d' < 0$ for all $(d, d') \in D$ (as defined in (11)). Define $g : X \times S \to \mathbb{R}$ by setting $g(y, s_1) = 0$ for all $y \in X$ and

$$g(y, s_i) \;:=\; \begin{cases} \sum_{j=1}^{i-1} -\hat{p}'_i & \text{if } y < x'; \\ \sum_{j=1}^{i-1} -\hat{p}_i & \text{otherwise,} \end{cases}$$

for all $i > 1$. The function g is supermodular; moreover, for any $\mu \in \Lambda(t)$, $\mu' \in \Lambda(t')$,

$$\int_S g(x', s)d\mu'(s) - \int_S g(x', s)d\lambda(s) + \int_S g(x, s)d\mu(s) - \int_S g(x, s)d\lambda'(s) < 0.$$

This in turn implies that $F(x,t) + F(x',t') < F(x,t') + F(x',t)$. **QED**

Continuation of Example 7. We show that correspondence Λ satisfies (F2). Take any $t' \geq t$ and $\lambda \in \Lambda(t)$, $\lambda' \in \Lambda(t')$. Given that $\int_S h(s)d(\lambda \wedge \lambda')(s) \leq \int_S h(s)d\lambda'(s) = t$ and $\int_S h(s)d\lambda'(s) = t'$, there is $\alpha \in [0,1]$ such that

$$\alpha \int_S h(s)d\lambda'(s) + (1-\alpha) \int_S h(s)d(\lambda \wedge \lambda')(s) \;=\; t.$$

Let $\mu = \alpha\lambda' + (1-\alpha)(\lambda \wedge \lambda')$ and $\mu' = \alpha\lambda + (1-\alpha)(\lambda \vee \lambda')$. Clearly, $\mu \in \Lambda(t)$, $\lambda' \succeq \mu$, and $\lambda \succeq \mu'$. Since $\lambda + \lambda' = (\lambda \vee \lambda') + (\lambda \wedge \lambda')$, we also obtain $\lambda + \lambda' = \mu + \mu'$. Hence,

$$\int_S h(s)d\mu'(s) \;=\; \int_S h(s)d\lambda(s) + \int_S h(s)d\lambda'(s) - \int_S h(s)d\mu(s) \;=\; t + t' - t \;=\; t'.$$

Thus $\mu' \in \Lambda(t')$. We conclude that Λ satisfies (F2). **QED**

Proof of Proposition 5. Proposition 4 guarantees that (i) implies (ii). To prove the converse, we first claim that if function f is supermodular for any supermodular function g that increases with s, then Λ satisfies the following property, which we shall refer to as (F3): for any $t' \geq t$, $\lambda \in \Lambda(t)$, and $\lambda' \in \Lambda(t')$, there is $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that $\lambda' \succeq \mu$ and $(1/2)\lambda + (1/2)\lambda' \succeq (1/2)\mu + (1/2)\mu'$. Notice that (F3) is weaker than (F2).

We prove by contradiction. Assuming that Λ violates (F3), we shall produce a function g that is supermodular in (x,s) and increasing in s such that f is *not* supermodular. Our proof is similar to the one we gave for the claim that (iii) implies (i) in Proposition 4 and we shall refer to it. Take any $t' \geq t$ and $\lambda \in \Lambda(t)$, $\lambda' \in \Lambda(t')$; suppose there is no $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ with the properties required by (F3). Then, by defining D as in (11), we obtain $\bar{C} \cap D = \emptyset$, where $\bar{C} = \left\{(c,c') \in \mathbb{R}^\ell \times \mathbb{R}^\ell : c \geq -c' \text{ and } c' \in \mathbb{R}^\ell_+\right\}$. Since D is compact and convex, and $\bar{C}$ is closed and convex with $(0,0) \in \bar{C}$, the strong separating hyperplane theorem guarantees that there is $(\hat{p}, \hat{p}')$ in $\mathbb{R}^\ell \times \mathbb{R}^\ell$ such that $\hat{p} \cdot d + \hat{p}' \cdot d < 0 \leq \hat{p} \cdot c + \hat{p}' \cdot c'$, for all $(c,c') \in \bar{C}$ and $(d,d') \in D$. Using the argument from the earlier proof we know that $\hat{p}' \geq \hat{p}$. Furthermore, in this case, $(c,0) \in \bar{C}$ for all $c \in \mathbb{R}^\ell_+$; therefore, $\hat{p} \geq 0$. We define $g : X \times S \to \mathbb{R}$ by (12). This function is supermodular since $\hat{p}' \geq \hat{p}$ and it increases with s since $\hat{p}'$, $\hat{p} \geq 0$. We have shown in the earlier proof that f in this case is not supermodular, yielding a contradiction.

To complete the proof we show that (F3) implies (F2) when Λ is upper comprehensive. (F3) states that for any $t' \geq t$, $\lambda \in \Lambda(t)$, and $\lambda' \in \Lambda(t')$, there is $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that $\mu(s_i) \geq \lambda'(s_i)$ and $\mu(s_i)+\mu'(s_i) \geq \lambda(s_i)+\lambda'(s_i)$ for all i. We modify μ and μ' such that the stronger property required by (F2) holds. This adjustment is done state by state, beginning with the lowest. Suppose $\mu(s_1)+\mu'(s_1) > \lambda(s_1)+\lambda'(s_1)$. If it is possible, choose $\nu^1(s_1)$ in the interval $\left[\lambda'(s_1), \mu(s_1)\right]$ such that $\nu^1(s_1)+\mu'(s_1) = \lambda(s_1)+\lambda'(s_1)$ and then set $\nu'^1(s_1) = \mu'(s_1)$. If, after setting $\nu^1(s_1) = \lambda'(s_1)$, we have $\nu^1(s_1)+\mu'(s_1) > \lambda(s_1)+\lambda'(s_1)$, then set $\nu'^1(s_1) = \lambda(s_1)$. Let $\nu^1(s_i) = \mu(s_i)$ and $\nu'^1(s_i) = \mu(s_i)$ for $i \geq 2$. Note that ν^1 and ν'^1 are bona fide distributions (i.e., both functions are increasing with the state) and, since Λ is upper comprehensive, $\nu^1 \in \Lambda(t)$, $\nu'^1 \in \Lambda(t')$. Furthermore, ν^1 and ν'^1 satisfy the conditions required by (F3) and $\nu^1(s_1) + \nu'^1(s_1) = \lambda(s_1) + \lambda'(s_1)$. Now define ν^2 and ν'^2 by $\nu^2(s_i) = \nu^1(s_i)$ and $\nu'^2(s_i) = \nu'^1(s_i)$, for all $i \neq 2$. If possible, set $\nu^2(s_2) \in \left[\max\{\lambda'(s_2), \nu^1(s_1)\}, \mu(s_2)\right]$ so that $\nu^2(s_1) + \nu'^1(s_2) = \lambda(s_2) + \lambda'(s_2)$ and then set $\nu'^2(s_2) = \nu'^1(s_2)$. If this is impossible, set $\nu^2(s_2) = \max\{\lambda'(s_2), \nu^1(s_1)\}$ and set $\nu'^2(s_2)$ so that $\nu^2(s_2) + \nu'^2(s_2) = \lambda(s_2) + \lambda'(s_2)$. Note that both ν^2 and ν'^2 are distributions, with $\nu^2 \in \Lambda(t)$, $\nu'^2 \in \Lambda(t')$, and $\nu(s_i) \geq \lambda'(s_i)$ for all i; furthermore, $\nu^2(s_i) + \nu'^2(s_i) \geq \lambda(s_i) + \lambda'(s_i)$ for all i, with equality in the case of $i = 1, 2$. Repeating this adjustment process we eventually obtain $\nu \in \Lambda(t)$ and $\nu' \in \Lambda(t')$ with the property that $\nu(s_i) \geq \lambda'(s_i)$ and $\nu(s_i) + \nu'(s_i) = \lambda(s_i) + \lambda'(s_i)$ for all i. Thus, (F2) holds. **QED**

Proof of Proposition 6. The proof is close to that of Proposition 4, so we shall only sketch it. To show that (i) $\Rightarrow$ (ii), take some $x' \geq x$, $t' \geq t$ and any $a \in \Gamma(x,t)$, $a' \in \Gamma(x',t')$. By definition of Γ, there are $\lambda, \lambda' \in \triangle_S$ such that $a_i = -\delta_i(x)\lambda(s_i)$, $a'_i = -\delta_i(x')\lambda'(s_i)$, for $i \leq \ell$, and $a_{\ell+1} = c(\lambda,t)$, $a'_{\ell+1} = c(\lambda',t')$. Property (C) guarantees that there are distributions $\mu, \mu' \in \triangle_S$ such that $\mu(s_i) - \lambda'(s_i) = \lambda(s_i) - \mu'(s_i) \geq 0$, for $i = 1, \ldots, \ell$, and $c(\lambda,t) + c(\lambda',t') \geq c(\mu,t) + c(\mu',t')$. Since g is supermodular, $\delta_i(x)$ is increasing in x, and we obtain $\delta_i(x')\left[\mu(s_i) - \lambda'(s_i)\right] \geq \delta_i(x)\left[\lambda(s_i) - \mu'(s_i)\right]$. Define vectors b, b' so that $b_i = -\delta_i(x')\mu(s_i)$, $b'_i = -\delta_i(x)\mu'(s_i)$, for $i \leq \ell$, and $b_{\ell+1} = c(\mu,t)$, $b'_{\ell+1} = c(\mu',t')$. Clearly, $b \in \Gamma(x',t)$, $b' \in \Gamma(x,t')$ and $a + a' \geq b + b'$.

To prove that (ii) $\Rightarrow$ (iii), note that $f(x,t) = g(x, s_{\ell+1}) + \min\left\{\mathbf{1} \cdot a : a \in \Gamma(x,t)\right\}$. An application of the Main Theorem guarantees that (iii) holds. .

To show that (iii) $\Rightarrow$ (i), suppose there is $t' \geq t$ and λ, λ' such that there is no

$\mu, \mu' \in \triangle_S$ satisfying the conditions required by (C). Then $D \cap K = \emptyset$, where

$$D = \Big\{(d, d', r) \in \mathbb{R}^\ell \times \mathbb{R}^\ell \times \mathbb{R} : d_i = \big[\mu'(s_i) - \lambda(s_i)\big], d'_i = \big[\mu(s_i) - \lambda'(s_i)\big], \\ \text{for } i = 1, 2, \ldots, \ell \text{ and } r \geq c(\mu, t) + c(\mu', t') - c(\lambda, t) - c(\lambda', t'), \text{ for } \mu, \mu' \in \triangle_S\Big\}$$

and $K = \big\{(-k, k, 0) \in \mathbb{R}^\ell \times \mathbb{R}^\ell \times \mathbb{R} : k \geq 0\big\}$. Both D and K are convex sets (the former because of the convexity of c) and $(D - K)$ is closed. By the strong hyperplane theorem to obtain $(\hat{p}, \hat{p}', q) \in \mathbb{R}^\ell \times \mathbb{R}^\ell \times \mathbb{R}$ such that $(\hat{p}, \hat{p}', q) \cdot (d, d', r) < 0 \leq (\hat{p}, \hat{p}') \cdot (-k, k)$, for all $(d, d', r) \in D$ and $(-k, k, 0) \in K$. In particular, $(\hat{p}, \hat{p}') \cdot (-k, k) \geq 0$ for all $k \geq 0$ guarantees that $\hat{p}' \geq \hat{p}$. If we choose $\mu = \lambda \wedge \lambda'$ and $\mu' = \lambda \vee \lambda'$, then $-d = d' > 0$, and so $(\hat{p}, \hat{p}', q) \cdot (d, d', r) = (\hat{p}, \hat{p}') \cdot (-d', d') + qr < 0$ implies that $q < 0$ (since $r > 0$). With no loss of generality, we may set $q = -1$. Take any $x, x' \in X$ such that $x' > x$ and define $g : X \times S \to \mathbb{R}$ as follows: let $g(y, s_1) = 0$, for all $y \in X$, and, for $i \geq 2$, let $g(y, s_i)$ have the form (12), which is supermodular. Moreover, for any μ, $\mu' \in \triangle_S$, we have

$$\begin{aligned}
&\left[\int_S g(x, s) d\lambda(s) + c(\lambda, t)\right] + \left[\int_S g(x', s) d\lambda'(s) + c(\lambda', t')\right] \\
&\quad - \left[\int_S g(x, s) d\mu'(s) + c(\mu', t')\right] - \left[\int_S g(x', s) d\mu(s) + (c(\mu, t)\right] \\
&\quad = \sum_{i=1}^{\ell} p_i \big[\mu'(s_i) - \lambda(s_i)\big] + \sum_{i=1}^{\ell} p'_i \big[\mu(s_i) - \lambda'(s_i)\big] \\
&\qquad - \big(c(\mu, t) + c(\mu', t') - c(\lambda, t) - c(\lambda', t')\big) \ < \ 0.
\end{aligned}$$

This leads to $f(x, t) + f(x', t') < f(x, t') + f(x', t)$, which contradicts (iii). **QED**

Proof of Proposition 7. It suffices to show that $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular in λ (for each t) and that it has decreasing differences in (λ, t). To prove the first claim, let λ, $\lambda' \in \triangle_S$ and denote $\lambda \vee \lambda'$ and $\lambda \wedge \lambda'$ by μ' and μ respectively. $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular in λ if, for all i,

$$d\lambda(s_i) \ln d\lambda(s_i) + d\lambda'(s_i) \ln d\lambda'(s_i) - \big[d\lambda(s_i) + d\lambda'(s_i)\big] \ln d\lambda^*(s_i, t) \\ \geq d(\mu)(s_i) \ln d\mu(s_i) + d\mu'(s_i) \ln d\mu'(s_i) - \big[d\mu(s_i) + d\mu'(s_i)\big] \ln d\lambda^*(s_i, t). \quad (13)$$

Clearly, this inequality holds with equality for $i = 1$. Consider $i > 1$. With no loss of generality, let $\mu(s_{i-1}) = \lambda(s_{i-1})$ and $\mu'(s_{i-1}) = \lambda'(s_{i-1})$. Consider two cases. Assume that (a) $d\lambda'(s_i) + \lambda'(s_{i-1}) \leq d\lambda(s_i) + \lambda(s_{i-1})$, so that $\mu(s_i) = \lambda(s_i)$ and $\mu'(s_i) = \lambda'(s_i)$. Then $d\mu(s_i) = d\lambda(s_i)$ and $d\mu'(s_i) = d\lambda'(s_i)$ and (13) is satisfied with equality. Suppose, instead, that (b) $d\lambda'(s_i) + \lambda'(s_{i-1}) > d\lambda(s_i) + \lambda(s_{i-1})$, which implies $\mu(s_i) = \lambda'(s_i)$ and $\mu'(s_i) = \lambda(s_i)$. Let $\delta = \lambda(s_{i-1}) - \lambda'(s_{i-1})$ and notice that $0 \leq \delta < d\lambda'(s_i) - d\lambda(s_i)$. Since $d\mu(s_i) = d\lambda'(s_i) - \delta$ and $d\mu'(s_i) = d\lambda(s_i) + \delta$,

$$\begin{aligned} d\mu(s_i) \ln d\mu(s_i) + d\mu'(s_i) \ln d\mu'(s_i) - \big[d\mu(s_i) + d\mu'(s_i)\big] \ln d\lambda^*(s_i, t) \\ = \big[d\lambda'(s_i) - \delta\big] \ln \big[d\lambda'(s_i) - \delta\big] + \big[d\lambda(s_i) + \delta\big] \ln \big[d\lambda(s_i) + \delta\big] \\ - \big[d\lambda(s_i) + d\lambda'(s_i)\big] \ln d\lambda^*(s_i, t) \\ \leq d\lambda(s_i) \ln d\lambda(s_i) + d\lambda'(s_i) \ln d\lambda'(s_i) - \big[d\lambda(s_i) + d\lambda'(s_i)\big] \ln d\lambda^*(s_i, t) \end{aligned}$$

where the last inequality follows from the convexity of the map from $z \to z \log z$. So we have shown that (13) holds for all i and thus $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular in λ.

In order to show that $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ has decreasing differences in (λ, t), take any distribution $\lambda' \succeq \lambda$, $t' \geq t$, and notice that

$$\begin{aligned} \Big[R\big(\lambda' \| \lambda^*(\cdot, t')\big) - R\big(\lambda \| \lambda^*(\cdot, t')\big)\Big] - \Big[R\big(\lambda' \| \lambda^*(\cdot, t)\big) - R\big(\lambda \| \lambda^*(\cdot, t)\big)\Big] \\ = \sum_{i=1}^{\ell} \big[\ln d\lambda^*(s_i, t') - \ln d\lambda^*(s_i, t)\big]\big[d\lambda(s_i) - d\lambda'(s_i)\big] \leq 0, \end{aligned}$$

since $\ln d\lambda^*(s, t') - \ln d\lambda^*(s, t)$ is increasing in i (because $\lambda^*(t)$ is increasing in t with respect to the monotone likelihood ratio order) and $\lambda' \succeq \lambda$. **QED**

Proof of Proposition 8. Let $v : X \times S \to \mathbb{R}$ be a continuous and bounded function. Since the problem is well-behaved we know that the function $(\mathscr{T}v)$, given by

$$\big(\mathscr{T}v\big)(x, s) = \max\Big\{F(x, s, y) + \beta(Av)(y, s) : y \in B(x, s)\Big\},$$

is a continuous function on $X \times S$ and $\mathscr{T}^n v$ converges uniformly to v^* as $n \to \infty$. By Proposition 4, whenever function v is supermodular, then so is Av. This implies that $F(x, s, y) + \beta(Av)(y, s)$ is supermodular over $X \times S \times X$. Given that the graph of correspondence B is a sublattice, by Theorem 4.3 in Topkis (1978), the function $\mathscr{T}v$ is supermodular in (x, s). Since supermodularity is preserved under uniform convergence, we conclude that $v^* = \mathscr{T}v^*$ is a supermodular function of (x, s). The set $\Phi(x, s)$ consists of elements y that maximize $F(x, s, y) + \beta(Av^*)(x, s)$ over $B(x, s)$. Since the objective function is supermodular, while values of correspondence B are complete sub-lattices of X, by the MCS theorem, set $\Phi(x, s)$ is a complete sub-lattice of X. Furthermore, since B increases over $X \times S$ in the strong set order, so does Φ. As the problem is well-behaved, $\Phi(x, s)$ admits the greatest selection $\phi(x, s)$ and this selection is increasing. That ϕ is Borel measurable follows from standard arguments (see HP). **QED**

References

Aliprantis, C. D., and K. C. Border (2006): *Infinite dimensional analysis: A hitchhiker's guide*. Berlin: Springer, 3rd edn.

Balbus, Ł., K. Reffett, and Ł. Woźny (2014): "A constructive study of Markov equilibria in stochastic games with strategic complementarities," *Journal of Economic Theory*, 150(C), 815–840.

Border, K. C. (1985): *Fixed point theorems with applications to economics and game theory*. Cambridge University Press.

Cherbonnier, F., and C. Gollier (2015): "Decreasing aversion under ambiguity," *Journal of Economic Theory*, 157(C), 606–623.

Curtat, L. (1996): "Markov equilibria of stochastic games with complementarities," *Games and Economic Behavior*, 17, 177–199.

Epstein, L. G., and M. Schneider (2003): "Recursive multiple-priors," *Journal of Economic Theory*, 113(1), 1–21.

——— (2010): "Ambiguity and asset markets," *Annual Review of Financial Economics*, 2(1), 315–346.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118(2), 133–173.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153.

Gollier, C. (2011): "Portfolio choices and asset prices: The comparative statics of ambiguity aversion," *Review of Economic Studies*, 78(4), 1329–1344.

Hopenhayn, H. A., and E. C. Prescott (1992): "Stochastic monotonicity and stationary distributions for dynamic economies," *Econometrica*, 60(6), 1387–1406.

Huggett, M. (2003): "When are comparative dynamics monotone?," *Review of Economic Dynamics*, 6(1), 1–11.

Iyengar, G. N. (2005): "Robust Dynamic Programming," *Mathematics of Operations Research*, 30(2), 257–280.

Jacobsen, S. E. (1970): "Production Correspondences," *Econometrica*, 38(5), 754–771.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity aversion, robustness, and the variational representation of preferences," *Econometrica*, 74(6), 1447–1498.

McFadden, D. (1966): "Cost, revenue, and profit functions: A cursory review," Working Paper 86, Institute of Business and Economic Research, University of California, Berkeley.

——— (1978): "Cost, Revenue, and Profit Functions," in *Production Economics: A Dual Approach to Theory and Applications*, ed. by M. Fuss, and D. McFadden, vol. 1, chap. 1. McMaster University Archive for the History of Economic Thought.

Milgrom, P., and J. Roberts (1990): "The economics of modern manufacturing: Technology, strategy, and organization," *American Economic Review*, 80(3), 511–28.

Milgrom, P., and C. Shannon (1994): "Monotone comparative statics," *Econometrica*, 62(1), 157–180.

Quah, J. K.-H. (2007): "The comparative statics of constrained optimization problems," *Econometrica*, 75(2), 401–431.

Sargent, T. J., and L. P. Hansen (2001): "Robust control and model uncertainty," *American Economic Review*, 91(2), 60–66.

Sharkey, W. (1981): "Convex games without side payments," *International Journal of Game Theory*, 10, 101–106.

Stachurski, J., and T. Kamihigashi (2014): "Stochastic stability in monotone economies," *Theoretical Economics*, 9(2), 383–407.

Stokey, N., R. Lucas, and E. Prescott (1989): *Recursive methods in economic dynamics*. Harvard University Press.

Strzalecki, T. (2011a): "Axiomatic foundations of multiplier preferences," *Econometrica*, 79(1), 47–73.

——— (2011b): "Probabilistic sophistication and variational preferences," *Journal of Economic Theory*, 146(5), 2117–2125.

Topkis, D. M. (1968): "Ordered optimal solutions," Ph.D. thesis, Stanford University.

——— (1978): "Minimazing a submodular function on a lattice," *Operations Research*, 26(2), 305–321.

——— (1998): *Supermodularity and complementarity*, Frontiers of economic research. Princeton University Press.