ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

Supermodular value functions and supermodular correspondences

Pawel Dziewulski and John K.-H.Quah

Number 795
First Published: April 2016
Revised: July 2017

Manor Road Building, Oxford OX1 3UQ

Supermodular value functions and supermodular correspondences*

Paweł Dziewulski† John K.-H. Quah‡

July 2017

Abstract

Economic problems often involve an objective function whose value depends partly on exogenous variables and partly on actions which are chosen by an agent. We are interested in the conditions under which the resulting value is a supermodular function of the exogenous variables (after actions have been optimally chosen). Problems of this sort arise in multi-output production, optimization with non-EU models, and dynamic programming. We show that these problems can be effectively tackled through a theory of supermodular correspondences. Our work builds on early results in monotone comparative statics in Topkis (1978), Hopenhayn and Prescott (1992), and Milgrom and Shannon (1994).

Keywords: monotone comparative statics, single crossing differences, multi-output production, ambiguity aversion, variational preferences, relative entropy, dynamic programming
JEL Classification: C61, D21, D24

1 Introduction

Consider a firm that uses n 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]

*We thank Eddie Dekel, Takashi Kamihigashi, and Tomasz Strzalecki for helpful comments.

†Department of Economics, University of Oxford. E-mail: pawel.dziewulski@economics.ox.ac.uk.

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

[1] For a proof of this result and extensions see Topkis (1978), Milgrom and Roberts (1990), and Milgrom and Shannon (1994).

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 could be represented by a correspondence Γ where, given a vector of factors x, set $\Gamma(x)$ consists of all the combinations of output goods that are producible using those factors. Assuming that there are m output goods priced at $q = (q_1, q_2, \ldots, q_m)$, factor-complementarity will hold if the maximum revenue

$$R(x) := \max \left\{ q \cdot y : y \in \Gamma(x) \right\}$$

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 the 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 $f(x, s)$, where $x \in X \subseteq \mathbb{R}$ is the chosen action at the state is $\tilde{s} \in S \subseteq \mathbb{R}$. The expected utility of action x is therefore $F(x, t) := \int f(x, \tilde{s}) d\lambda(\tilde{s}, t)$, where $t \in T \subseteq \mathbb{R}$ parametrises the distribution function $\lambda(\cdot, t)$ over S. Suppose that f is such that the marginal payoff of a higher action increases with s, i.e., function f is supermodular. Then it seems reasonable that the *expected* marginal payoff of a higher action should be greater when higher states are more likely. This intuition is correct: if f is a supermodular function of (x, s), then F is a supermodular function of (x, t) if $\lambda(\cdot, t)$ is increasing with t in the following sense: if $t' \geq t$ then $\lambda(\cdot, t')$ first order stochastically dominates $\lambda(\cdot, t)$.

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

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

where $\Lambda(t)$ denotes a set of distributions over S. Note that G is the value function arising from Nature choosing $\lambda \in \Lambda(t)$. Assuming that f is supermodular, what conditions on the correspondence Λ will guarantee that G is supermodular?

Our final example is drawn from dynamic programming, where it is often convenient to establish that the value function is supermodular in terms of the exogenous and endogenous state variables. Suppose that at each period, the agent chooses an action $y \in X \subseteq \mathbb{R}$, after observing the current exogenous state $s \in S \subseteq \mathbb{R}$, which is random.

[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 3.

The instantaneous payoff is $u(x, s, y)$ where $x \in X$ is the action chosen in the previous period. By definition, the value $v^*(x, s)$ is the maximal expected discounted utility of the agent, conditional on x being the previous period's action and s being the current exogenous state. Following standard arguments (see Stokey, Lucas, and Prescott, 1989), v^* is the fixed point of the operator $\mathscr{T}$ that maps a real-valued function $w : X \times S \to \mathbb{R}$ to the function $(\mathscr{T}w) : X \times S \to \mathbb{R}$, given by

$$(\mathscr{T}w)(x, s) = \max \left\{ u(x, s, y) + \beta \int_S w(y, \tilde{s}) d\lambda(\tilde{s}, s) : y \in X \right\}, \tag{1}$$

where $\lambda(\cdot, s)$ is the distribution over S in the following period, conditional on s being the exogenous state in the current period. To guarantee that v^* is supermodular, a crucial step in the argument is to show that $\mathscr{T}$ maps a supermodular function w into another supermodular function: once again, the nub of the issue concerns the supermodularity of a value function. In their monotone theory of dynamic programming, Hopenhayn and Prescott (1992) make use of the following result of Topkis (1978), which we shall refer to as the *Topkis Value Theorem*: if $h(\theta, z)$ is jointly supermodular in (θ, z), then $H(\theta) := \max_{z \in Z} h(\theta, z)$ is a supermodular function of θ.[3] It is straightforward to show that the objective function on the right hand side of (1) will be supermodular in (x, s, y) if u is supermodular, $\lambda(\cdot, s)$ is increasing in s with respect to first order stochastic dominance order, and w is supermodular; it then follows from the Topkis Value Theorem that function $(\mathscr{T}w)$ is supermodular in (x, s).

But what if u is not supermodular? Could we guarantee the supermodularity of the value function by appealing to an extended version of the Topkis Value Theorem? We know that ordinal versions of supermodularity, as developed in Milgrom and Shannon (1994), are sufficient to guarantee that $\arg\max_{z \in Z} h(\theta, z)$ is increasing in θ. Could these conditions be used to extend the Topkis Value Theorem?

Our results. In Sections 2 and 3, we introduce a generalized notion of supermodularity for a correspondence $\Gamma : X \to Y$, where X is a lattice and Y an ordered vector space. We show that it is sufficient (and, in a sense, necessary) to guarantee that the function mapping x to $\max \{\phi(y) : y \in \Gamma(x)\}$ is supermodular, for any positive linear functional $\phi : Y \to \mathbb{R}$. Similarly, we develop a related notion of supermodularity of Γ which guarantees that the map from x to $\min \{\phi(y) : y \in \Gamma(x)\}$ is supermodular.

[3] To keep it simple, think of θ and z as elements of Euclidean spaces Θ and Z.

Section 4 is devoted to exploiting these results in different contexts. Their implications for production analysis are found in Applications 1, 2, and 3. We formulate a notion of *generalized complementarity* in production that is applicable to any subset of goods (which can be input goods, output goods, or a combination of both) and show what properties of the production technology are sufficient to guarantee complementarity within that set. We also identify the conditions under which a firm raises its output when the price of an input falls. Lastly, we ask when a firm will have a supermodular production function if it could switch among different technologies.

Our extension of the Topkis Value Theorem can be found in Application 3; we weaken the joint supermodularity condition on the objective function by partially replacing it with versions of the single crossing property as developed in Milgrom and Shannon (1994).

Applications 4 and 5 deal with the comparative statics of decision-making with maxmin, variational, or multiplier preferences. For each of these preference models, we formulate what it means for "*beliefs to shift towards higher states*" in a way that leads to an agent choosing a higher action whenever that occurs.

We consider applications to dynamic programming in Section 5; specifically we extend the monotone method of Hopenhayn and Prescott (1992) in two directions. In Application 6, we show how the supermodularity of the instantaneous payoff function can sometimes be weakened to a property we call *conditional increasing differences*; this result makes use of our extension of the Topkis Value Theorem (in Application 3). Second, we show in Application 7 that Hopenhayn and Prescott's method can be extended to the case where, instead of maximizing expected discounted utility, the agent's preference over uncertain future utility streams conforms to the maxmin model.

2 Basic concepts

A binary relation on a set X is a *partial order* on X if it is reflexive, transitive and antisymmetric.[4] A *partially ordered set*, or simply a *poset*, is a pair $(X, \geq_X)$ consisting of a set X and a partial order $\geq_X$ on X. We denote the strict counterpart of $\geq_X$ by $>_X$, that is, for any x and x' in X, we have $x' >_X x$, if $x' \geq_X x$ and $x' \not\leq_X x$. Whenever it causes no confusion, we abbreviate our notation by denoting $(X, \geq_X)$ with X.

[4] See Topkis (1998) for a detailed introduction to the concepts covered here.

For any two elements x, x' of a poset X, their *least upper bound*, or the *join*, is denoted by $x \vee x'$, while their *greatest lower bound*, or the *meet*, by $x \wedge x'$, where both elements are defined with respect to the partial order $\geq_X$. A poset X is a *lattice* if, for any x, x' in X, both their join $x \vee x'$ and their meet $x \wedge x'$ belong to the set. A poset Y is a *sublattice* of X if it is a subset of X that contains elements $y \vee y'$, $y \wedge y'$, for any y, y' in Y.

For the purposes of this paper, the most important lattice is the Euclidean space $(\mathbb{R}^\ell, \geq)$, where we say that $x \geq y$ if $x_i \geq y_i$ for all $i = 1, 2, \ldots, \ell$. In this case, for vectors x and y in $\mathbb{R}^\ell$, we have $(x \vee y)_i = \max\{x_i, y_i\}$ and $(x \wedge y)_i = \min\{x_i, y_i\}$.

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

A binary relation on a set X is a *preorder* if it is reflexive and transitive. Our generalisation of supermodularity applies to correspondences that map a lattice to what we shall call an *ordered vector space*; this refers to a real vector space endowed with a preorder that is preserved by the vector space operations. In other words $(Y, \geq_Y)$ is an ordered vector space whenever Y is a vector space and $\geq_Y$ is a preorder satisfying the following properties: if $x \geq_Y y$ then $x + z \geq_Y y + z$ and $\alpha x \geq \alpha y$, for any x, y, z in Y and $\alpha \geq 0$. Clearly, the Euclidean space is an ordered vector space. Another important example is the space of signed finite measures defined on a partially ordered measurable space $(S, \mathcal{S})$. This is a real vector space which contains, crucially for our purposes, the set of probability measures. Since S is partially ordered, the signed measures can be ranked with respect to first order stochastic dominance, i.e., for any measures μ and ν in Y, we have $\mu \geq_Y \nu$ if $\int_S f d\mu \geq \int_S f d\nu$, for any bounded and measurable function $f : S \to \mathbb{R}$ that is increasing on S with respect to the corresponding partial order.

2.1 Upper and lower supermodularity

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

$$z + z' \;\geq_Y\; y + y'. \tag{2}$$

Equivalently, the vectors need to satisfy $(z + z')/2 \geq_Y (y + y')/2$. See Figure 1 for a graphical interpretation of upper supermodularity. A correspondence $\Gamma : X \to Y$ is *lower*

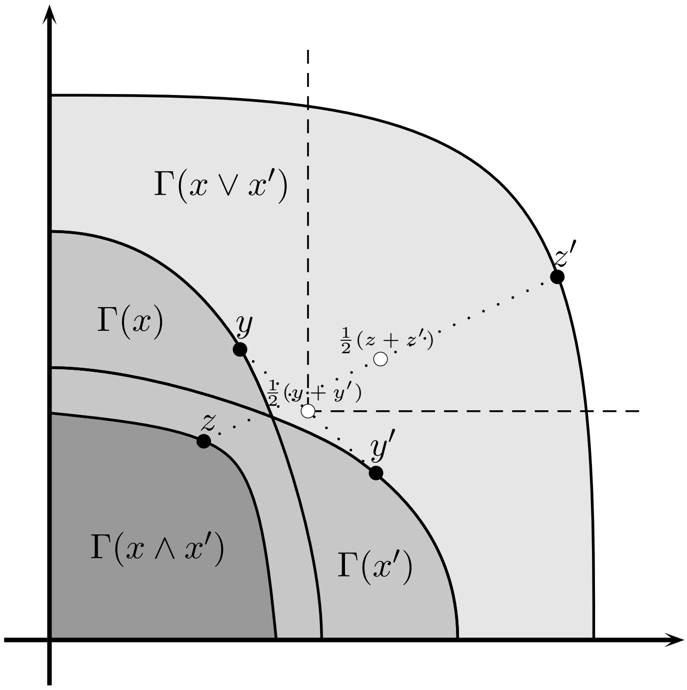

Figure 1: An upper supermodular correspondence for $Y = \mathbb{R}^2_+$.

supermodular if for any two elements x, x' in X and $z \in \Gamma(x \wedge x')$, $z' \in \Gamma(x \vee x')$ there are some vectors $y \in \Gamma(x)$ and $y' \in \Gamma(x')$ that satisfy condition (2).[5] Finally, we say that the correspondence is *supermodular* once it is both upper and lower supermodular.

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 $z \in \Gamma(x \wedge x')$ and $z' \in \Gamma(x \vee x')$ such that (2) holds with the inequality reversed; equivalently, Γ is submodular if $-\Gamma$ is upper supermodular. Similarly, one may define lower submodularity and submodularity.

Clearly, our definition of supermodular correspondences generalises the familiar notion of supermodularity applied to real-valued functions, introduced at the beginning of this section. This notion also extends the concept of *stochastic supermodularity* introduced in Topkis (1968) to correspondences;[6] a function mapping a lattice to the set of probability measures on some measurable space is said to be stochastically supermodular if condition (2) holds with $\geq_Y$ representing first order stochastic dominance.

Suppose that $\Gamma : X \to Y$ has *downward comprehensive* values, i.e., $y \in \Gamma(x)$ and

[5] Notice that, the distinction between upper and lower supermodularity disappears if Γ is a function, i.e., Γ is singleton-valued, rather than a set-valued correspondence.

[6] Though Topkis (1968) refers to this property as *stochastic convexity*, the term *stochastic supermodularity* is more commonly used; see Curtat (1996) or Balbus, Reffett, and Woźny (2014) for applications.

$y \geq_Y y'$ implies $y' \in \Gamma(x)$, for any y, y' in Y and x in X. In this case, it is straightforward to show that Γ is upper supermodular if and only if

$$\Gamma(x \wedge x') + \Gamma(x \vee x') \supseteq \Gamma(x) + \Gamma(x'), \quad \text{for all } x,\, x' \in X. \tag{3}$$

The fact that (3) implies upper supermodularity is clear and does not even require the downward comprehensiveness of Γ. To show that the converse holds, note that if Γ is upper supermodular, then for any $y \in \Gamma(x)$ and $y' \in \Gamma(x')$ there is some $z \in \Gamma(x \wedge x')$ and $z' \in \Gamma(x \vee x')$ such that $z + z' \geq_Y y + y'$. Hence, $z \geq_Y (y + y' - z')$ and, since Γ is downward comprehensive, we have $(y + y' - z) \in \Gamma(x \wedge x')$. Consequently, this implies that $(y + y') = (y + y' - z) + z$ is an element of $\Gamma(x \wedge x') + \Gamma(x \vee x')$.

A special case of property (3) appears in the study of cooperative games with non-transferable utility. In that context, set X is interpreted as the collection of non-empty coalitions of a finite set N of players in a game, endowed with is the set inclusion order $\geq_X$, which makes X 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 (3) holds (see Sharkey, 1981, Section 2).

2.2 Examples of supermodular correspondences

The following is a list of simple ways to construct supermodular correspondences.

Example 1. Suppose that function $f_i : X \to \mathbb{R}$ is supermodular over a lattice X, for all $i = 1, \ldots, \ell$. The map $F : X \to \mathbb{R}^\ell$, given by $F(x) := \big(f_1(x), \ldots, f_n(x)\big)$ is a supermodular function, i.e., we have $F(x \wedge x') + F(x \vee x') \geq F(x) + F(x')$, for all x, x' in X, where $\geq$ denotes the coordinate-wise partial order on $\mathbb{R}^\ell$.

Example 2. Consider correspondence $\Gamma_i : X_i \to Y$, where $X_i \subseteq \mathbb{R}$ and Y is an ordered vector space, for $i = 1, 2$. 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).

Example 3. For any subset Z of an ordered vector space Y, a supermodular function $f : X \to Y$ over a lattice X, and positive scalars α and β, the mapping $\Gamma : X \to Y$, given by $\Gamma(x) := \big\{\alpha y + \beta f(x) : y \in Z\big\}$, is a supermodular correspondence.

Suppose that Y is the space of finite signed measures endowed with the first order stochastic dominance order. Whenever $Z \subseteq Y$ is a set of probability measures, while $f(x)$

is a probability measure for all $x \in X$, then for any scalars α and β such that $\alpha + \beta = 1$, set $\Gamma(x)$ is a subset of probability measures. This provides an easy way of constructing supermodular correspondences that map a lattice to the space of probability measures.

Example 4. Let Z be a convex subset of an ordered vector space Y, such that $z \geq 0$, for all $z \in Z$. For any supermodular function $f : X \to \mathbb{R}_+$ over a lattice X, the correspondence $\Gamma : X \to Y$ given by $\Gamma(x) := \{f(x)z : z \in Z\}$ is supermodular.

This claim requires a short proof. Since Z is convex and non-negative, Lemma 5.27 in Aliprantis and Border (2006) guarantees that $\alpha Z + \beta Z = (\alpha + \beta)Z$, for any positive scalars α and β. To show that Γ is upper supermodular, take any $f(x)y \in \Gamma(x)$ and $f(x')y' \in \Gamma(x')$. Given the above property of set Z, there is some vector $v \in Z$ such that $f(x)y + f(x')y' = [f(x) + f(x')]v$. Moreover, by supermodularity of function f,

$$[f(x) + f(x')]v \leq [f(x \wedge x') + f(x \vee x')]v.$$

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

Example 5. Let X and T be lattices and Z a sublattice of $X \times T$. We denote by X_Z the set of elements in X for which there is t such that $(x, t) \in Z$; for each $x \in X_Z$, the sectional set $Z(x) = \{t \in T : (x, t) \in Z\}$ is non-empty. It is straightforward to check that X_Z is a sublattice of X and for each $x \in X_s$, the set $Z(x)$ is a non-empty sublattice of T. Suppose that $f : Z \to Y$ is a supermodular function, where Y is an ordered real vector space. Then the correspondence $\Gamma : X_Z \to Y$, given by

$$\Gamma(x) := \{f(x, t) : t \in Z(x)\},$$

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 $y = f(x, t)$ and $y' = f(x', t')$. Moreover, the supermodularity of function f says that

$$f\big((x \wedge x'), (t \wedge t')\big) + f\big((x \vee x'), (t \vee t')\big) \geq f(x, t) + f(x', t').$$

Hence, the element $f\big((x, t) \wedge (x', t')\big)$ belongs to $\Gamma(x \wedge x')$ and $f((x, t) \vee (x', t'))$ is in $\Gamma(x \vee x')$, which concludes our argument.

It is straightforward to show that upper supermodularity is preserved by downward comprehensive transformations of correspondences. That is, for any upper supermodular correspondence $\Gamma : X \to Y$, mapping $\Lambda(x) := \{y \in Y : y \leq_Y z, \text{ for some } z \in \Gamma(x)\}$ is an upper supermodular correspondence. Analogously, lower supermodularity is preserved by upward comprehensive transformations.

Last but not least, it is clear that both *upper and lower supermodularity are preserved by weighted sums*, i.e., for any upper (lower) supermodular correspondences $\Gamma, \Lambda : X \to Y$, mapping $\Omega(x) := \alpha\Gamma(x) + \beta\Lambda(x)$ is an upper (lower) supermodular correspondence, for any positive scalars α and β. It follows that, when given examples of supermodular correspondences, we could create even more by simply adding them up.

3 Value functions of supermodular correspondences

In this section we present our main results 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 $\phi : Y \to \mathbb{R}$,*[7]

(i) *if correspondence $\Gamma : X \to Y$ is upper supermodular then the function $f : X \to \mathbb{R}$, given by $f(x) := \max\{\phi(y) : y \in \Gamma(x)\}$, is supermodular;*[8]

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

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

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

[7] 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.

[8] 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).

where the first inequality follows from ϕ being positive, while the second is implied by the definition of supremum. By taking the supremum over the left side of the inequality, we conclude that $f(x)+f(x') \leq f(x\wedge x')+f(x\vee x')$. Hence, function f is supermodular.

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

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

where the first inequality follows from ϕ being positive and the second is implied by the definition of supremum. Once we take the infimum on the left of this inequality, we obtain $f(x\wedge x')+f(x\vee x') \geq f(x)+f(x')$, which concludes the proof. □

In some applications one would like to investigate submodular properties of the value functions; in those instances the following analogue to the Main Theorem may apply. We shall skip the proof since it is similar to the one for the Main Theorem.

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 $f(x) := \min\{\phi(y) : y \in \Gamma(x)\}$, is submodular;*

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

We now give a flavour of how the Main Theorem can be applied to the analysis of firm behaviour. A fuller discussion is provided in Section 4. Suppose that $Y = \mathbb{R}^J$; in this case, any positive linear functional on Y is of the form $\phi(y) = q\cdot y$, where q is a vector in $\mathbb{R}^J_+$. Part (i) of the Main Theorem tells us that, so long as $q \geq 0$, then

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

is a supermodular function if Γ is upper supermodular. Recall the first motivating example we considered in the Introduction, concerning a firm producing J different output goods using I inputs. In that case, the correspondence $\Gamma : \mathbb{R}^I_+ \to \mathbb{R}^J$, where $\Gamma(x) \subseteq \mathbb{R}^J_+$, gives all combinations of output goods that could be produced using the input vector x. If $q \in \mathbb{R}^J_+$ are the output prices, then $f(x)$ is the maximum revenue obtainable by the firm, given the employment of x. A related but slightly different interpretation of f is to suppose that the firm is operating in a risky environment with J states of the world. Then the set $\Gamma(x) \subseteq \mathbb{R}^J_+$ gives 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 $f(x)$ is the greatest expected revenue achievable under x. Regardless of the interpretation, the Main Theorem guarantees that *function f is supermodular whenever the production correspondence Γ is upper supermodular.*

Example A. For a specific example of an upper supermodular output correspondence, suppose there are three inputs and two outputs (or state contingent revenues), where

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

In the above example, input 1 is non-rivalrous in the sense that 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.

To see that the above correspondence is upper supermodular, first notice that set

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

is a sublattice of $\mathbb{R}^4$. Moreover, $h : Z \to \mathbb{R}^2$, where $h(x,t) := \big(\sqrt[3]{x_1\, x_2\, t}, \sqrt{x_1} + \sqrt{x_3 - t}\big)$, is a supermodular function. Therefore, by the claim made in Example 5, the correspondence $\Lambda(x) := \big\{h(x,t) : (x,t) \in Z\big\}$ is upper supermodular. As Γ is a downward comprehensive transformation of Λ, it is also upper supermodular.

It is not hard to see that the assumptions in the Main Theorem are essentially tight. The following result (which we prove in the Appendix) gives a converse to the theorem in the case where Y is an Euclidean space. The assumption that Γ is convex- and compact-valued is standard, at least when we interpret Γ as a production correspondence.

Proposition 1. *Suppose X is a lattice and Y is an Euclidean space. Moreover, let correspondence $\Gamma : X \to Y$ have compact and convex values.*

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

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

4 Applications

The results developed in the last section lead to a wide range of applications that confirm the value of extending the concept of supermodularity to correspondences. Before we discuss these applications, it is useful to give a quick survey of some basic concepts and tools in monotone comparative statics.

For any two subsets Y, Y' of a lattice $(X, \geq_X)$, 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 $(y \vee y') \in Y'$. In particular, when $Y := \{y\}$ and $Y' := \{y'\}$, then Y' dominates Y in this sense if and only if $y' \geq_X y$. Whenever Y and Y' both contain their greatest elements, denoted by y and y', respectively, then $y' \geq_X y$ if Y' dominates Y in the strong set order. Similarly, if Y and Y' contain their least elements z and z' respectively, then $z' \geq_X z$. A subset Y of X is a sublattice if and only if Y dominates itself in the strong set order. While the strong set order is obviously not complete, it is transitive over the subsets of X (see Topkis, 1978). The basic results outlined below provide conditions under which the set of maximizers of some objective function is increasing in the strong set order.

Cardinal conditions for monotone comparative statics. Let X be a lattice and T be a partially order 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 $f(x', t) - f(x, t)$ is increasing in t. Note that this notion is closely related to supermodularity; indeed if T is totally ordered (and hence a lattice), then it is straightforward to check that 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 set $\arg\max_{x \in X} f(x, t)$ is a sublattice of X if $f : X \times T \to \mathbb{R}$ is supermodular in x. Moreover, if in addition f has increasing differences in (x, t), then $\arg\max_{x \in X} f(x, t)$ increases with t in the strong set order, i.e., $\arg\max_{x \in X} f(x, t')$ dominates $\arg\max_{x \in X} f(x, t)$ in the strong set order for any $t' \geq_T t$. We shall refer to this result as the *Topkis Monotone Comparative Statics (MCS) Theorem*. We also know that when $\arg\max_{x \in X} f(x, t)$ is a *compact* sublattice of a Euclidean space, then the greatest and least element of $\arg\max_{x \in X} f(x, t)$ exist and are increasing with t.

As a basic application of this result, suppose a firm is endowed with a supermodular production function $f : \mathbb{R}^I_+ \to \mathbb{R}_+$ and can sell its output at the price 1. Given input prices

$p = (p_1, \ldots, p_I)$ in $\mathbb{R}^I_{++}$ and an input profile x, the firm's profit is $\pi(x, p) = f(x) - p \cdot x$. It is straightforward to verify that π has increasing differences in $(x, -p)$, and that it is a supermodular function of x if f is a supermodular function. It then follows from the Topkis MCS Theorem that $\mathcal{M}(p) := \arg\max_{x \in \mathbb{R}^I_+} \pi(x, p)$ is increasing in $-p$ in the strong set order. In this sense, the factors of production are *complements* since the demand for *all* factors increase (weakly) when the price of any factor i falls.

Ordinal conditions for monotone comparative statics. Milgrom and Shannon (1994) point out that, for the purposes of comparative statics, the cardinal conditions in Topkis's result are plainly unnecessary. They develop ordinal versions of those conditions and show that in many significant economic applications, this distinction is important. The function $f : X \times T \to \mathbb{R}$ is *quasisupermodular* in x if, for any $t \in T$, $f(x, t) \geq (>)\ f(x \wedge x', t)$ implies $f(x \vee x', t) \geq (>)\ f(x', t)$. The function f has *single crossing differences in* (x, t) whenever $f(x', t) \geq (>)\ f(x, t)$ implies $f(x', t') \geq (>)\ f(x, t')$, for any $x' \geq_X x$ and $t' \geq_T t$.[9] Note that single crossing differences in (x, t) is not the same as single crossing differences in (t, x), even when X and T are both lattices. It is clear that both quasi-supermodularity and single crossing differences are ordinal properties, i.e., they are preserved by increasing transformations, and are respectively weaker than supermodularity and increasing differences on f. The *Milgrom-Shannon MCS Theorem* says that $\arg\max_{x \in X} f(x, t)$ is a sublattice of X whenever f is quasi-supermodular and if, in addition, f obeys single crossing differences in (x, t), then $\arg\max_{x \in X} f(x, t)$ is increasing with t. In many applications, X and T are both totally ordered, e.g., subsets of the real line; in this case, this result says that the set of optimal solutions is monotone so long as f obeys single crossing differences in (x, t).

Application 1: Generalised complementarity in production

Consider a firm endowed with a technology that employs I inputs to manufacture J output goods. We represent this technology by a production possibility set $P \subseteq \mathbb{R}^I_+ \times \mathbb{R}^J_+$. Abusing the notation, we denote the disjoint sets of inputs and outputs by I and J, respectively. An arbitrary element z of P denotes a feasible production profile that uses

[9] Milgrom and Shannon (1994) refers to this as the *single crossing property*. The term *single crossing differences*, follows Milgrom (2004).

$(z_k)_{k\in I}$ units of input to produce an output profile $(z_k)_{k\in J}$.[10] Conditional on prices $p_k > 0$, for each good $k \in I \cup J$, the firm chooses $z \in P$ in order to maximise its profit, which is given by $\sum_{k\in J} p_k z_k - \sum_{k\in I} p_k z_k$.

In Section 3, we provided conditions under which the firm's revenue is a supermodular function of inputs, which ensures that all input goods are complements. We now consider the problem of guaranteeing complementarity among an arbitrary subset C of commodities in $I \cup J$, which may consist of both input and output goods. Since (as far as we know) this issue has not been studied before, we first need to provide a definition that captures our intuition of what complementarity could mean in this context.

Generalized Complementarity Given a subset C of $I \cup J$, denote its complement by $C' := (I \cup J) \setminus C$. Let $X := \{(z_k)_{k\in C} : z \in P\}$, which is a subset of $\mathbb{R}^C_+$.[11] We say that the *goods in* C *are complements* whenever the correspondence $\mathcal{M} : \mathbb{R}^{I+J}_{++} \to X$, given by

$$\mathcal{M}\big((p_k)_{k\in C'}, (p_k)_{k\in C\cap J}, (p_k)_{k\in C\cap I}\big) = \Big\{ (z_k)_{k\in C} \;:\; z' \in \arg\max_{z\in P} \sum_{k\in J} p_k z_k - \sum_{k\in I} p_k z_k \Big\}, \quad (4)$$

increases in the strong set order with respect to $\big((p_k)_{k\in C\cap J}, (-p_k)_{k\in C\cap I}\big)$, keeping all other prices fixed.[12] This definition captures the idea that, whenever there is (i) an increase of the price of an output in C or (ii) a fall in the price of an input good in C, then *all* goods in C will be 'favourably' affected, in the sense that both the firm's demand for input goods in C, as well as its production of output goods in C, will increase.

To formulate a sufficient condition for goods in C to be complements, we first define the production possibility correspondence $\Gamma : X \to \mathbb{R}^{C'}$, that maps vectors $x \in X \subset \mathbb{R}^C_+$ to those combinations of goods in C' that are feasible given the firm's technology, with the proviso that inputs in C' enter with the negative sign. To be precise, let

$$\Gamma(x) := \Big\{ y \in \mathbb{R}^{C'} : \text{there is some } z \in P \text{ such that } z_k = x_k, \text{ for all } k \in C, \\ \text{while } z_k = y_k, \text{ for all } k \in C' \cap J, \text{ and } z_k = -y_k, \text{ for all } k \in C' \cap I \Big\}. \quad (5)$$

[10] A reader may notice that our definition of a production set is not the usual one, because we have not adopted the convention of writing inputs as negative entries in a production vector. The formulation we adopt is more convenient for our purpose.

[11] We abuse notation by denoting the sets C and C' and their cardinalities in the same manner.

[12] To keep the exposition simple, we are assuming that there is a solution to the firm's profit-maximization problem at all price vectors.

Proposition 2. *The goods in C are complements if set X is a sublattice of $\mathbb{R}^C_+$ and correspondence $\Gamma : X \to \mathbb{R}^{C'}$, defined in (5), is upper supermodular.*

Proof. The firm's profit maximisation problem can be formulated as a two step procedure. First, for each $x \in X$, we determine the maximal revenue achievable given the technology; assuming that the price of good $k \in C'$ is p_k, the maximal revenue is given by function $f(x) := \max\big\{ \sum_{k\in C'} p_k y_k : y \in \Gamma(x)\big\}$. The firm chooses $x \in X$ to maximise its profits

$$\Pi\big(x, (p_k)_{j\in C\cap J}, (p_k)_{k\in C\cap I}\big) \;=\; f(x) - \sum\nolimits_{k\in C\cap I} p_k x_k + \sum\nolimits_{k\in C\cap J} p_k x_k.$$

Notice that, given the definition of $\mathcal{M}$ in (4),

$$\arg\max_{x\in X} \Pi\big(x, (p_k)_{k\in C\cap J}, (p_k)_{k\in C\cap I}\big) = \mathcal{M}\big((p_k)_{k\in C'}, (p_k)_{k\in C\cap J}, (p_k)_{k\in C\cap I}\big).$$

From this observation, and from the Topkis MCS Theorem, we know that the goods in C are complements if Π is a supermodular function and has in increasing differences in $\big(x, \big((p_k)_{k\in C\cap J}, (-p_k)_{k\in C\cap I}\big)\big)$. The latter property is always true given the formula for Π, while Π is supermodular in x if f is supermodular. The Main Theorem guarantees that f is supermodular if the correspondence Γ is upper supermodular. □

The following examples show applications of this proposition.

Example B. We are interested in conditions under which all outputs are complements (i.e., $C = J$) when the firm's production possibility set is given by

$$P := \big\{(y, x) \in \mathbb{R}^I_+ \times \mathbb{R}^J_+ : g(y) \geq h(x)\big\},$$

where $g : \mathbb{R}^I_+ \to \mathbb{R}_+$ and $h : \mathbb{R}^J_+ \to \mathbb{R}_+$ are strictly increasing functions.[13] In this case, for each output vector x in $X = \mathbb{R}^J_+$, we have

$$\Gamma(x) := \big\{y \in \mathbb{R}^I_- : g(-y) \geq h(x)\big\}.$$

We claim that, whenever function h is submodular and g is concave and homogeneous of degree 1, then Γ is supermodular, and thus upper supermodular, so that Proposition 2 applies. Indeed, define the set $Z := \big\{z \in \mathbb{R}^I_+ : g(z) \geq 1\big\}$, which is positive and convex. By our claim in Example 4, the correspondence $\Lambda(x) := -h(x)Z$ is supermodular. Furthermore, the homogeneity of function g guarantees that $\Lambda(x) = \Gamma(x)$.

[13] We could interpret $h(x)$ as the level of some intermediate good which can be produced with x; this intermediate good can then be transformed into different output goods via the function g.

Example C. Consider a firm producing outputs a and b using capital k and labour ℓ. The use of capital is non-rivalrous but each unit of labour can be assigned to the production of either a or b, but not both. Suppose that input (k, ℓ_a) allows to produce up to $g(k, \ell_a)$ units of output a, while $h(k, \ell_b)$ is the output of good b when (k, ℓ_b) is employed. Hence, the firm's production possibility set is given by

$$P := \Big\{(k, \ell, a, b) \in \mathbb{R}^4_+ : \; a \leq g(k, \ell_a), \, b \leq h(k, \ell_b), \text{ with } \ell_a + \ell_b = \ell\Big\}.$$

We claim that *capital k and output a are complements* when the following conditions hold: (i) $h : \mathbb{R}^2_+ \to \mathbb{R}_+$ is a supermodular function; (ii) $g : \mathbb{R}^2_+ \to \mathbb{R}_+$ is continuous, strictly increasing, and supermodular, while $g(k, \cdot)$ is unbounded and concave, for all $k > 0$; and (iii) both factors are essential, i.e., $g(k, 0) = g(0, \ell_a) = 0$ for all $k \geq 0$ and $\ell_a \geq 0$.

Let $X := \big\{(k, a) \in \mathbb{R}^2 : \; (k, \ell, a, b) \in P \text{ for some } \ell \text{ and } b\big\}$. Since g is continuous, $g(k, \cdot)$ is unbounded, while $g(k, 0) = 0$, for any $k > 0$ and $a \geq 0$, there is a unique $\phi(k, a) \geq 0$ such that $g\big(k, \phi(k, a)\big) = a$. In other words, $\phi(k, a)$ is the least amount of labour needed to produce a when k units of capital are used. As ϕ is well defined for all $k > 0$ and $a \geq 0$, while $g(k, 0) = g(0, \ell_a) = 0$, for all $k \geq 0$ and $\ell_a \geq 0$, this implies that $X = (\mathbb{R}_{++} \times \mathbb{R}_+) \cup \big\{(0, 0)\big\}$, which is a lattice. For any $(k, a) \in X$, let

$$\Gamma(k, a) = \Big\{(-\ell, b) \in \mathbb{R}^2 : -\ell \leq -\ell_b - \phi(k, a) \text{ and } b \leq h(k, \ell_b), \; \text{for any } \ell_b \geq 0\Big\}.$$

It is easy to check that since g is monotone, supermodular, and concave in ℓ, the function ϕ is submodular.[14] This implies that $H(k, a, \ell_b) := \big(-\ell_b - \phi(k, a), h(k, \ell_b)\big)$ is a supermodular function over a sublattice of $\mathbb{R}^3$. Given the claim in Example 5, correspondence $\Lambda(k, a) := \big\{H(k, a, \ell_b) : \ell_b \geq 0\big\}$ is upper supermodular. Since Γ is a downward comprehensive transformation of Λ, it is upper supermodular, and so Proposition 2 applies.

Application 2: Factor prices and output[15]

Suppose a firm produces a single output using I inputs and has the production function $f : \mathbb{R}^I_+ \to \mathbb{R}_+$. We assume that the firm derives some benefit from output q, which we denote by $B(q)$ in $\mathbb{R}$. The objective of the firm is to choose inputs x in order to maximise $B\big(f(x)\big) - p \cdot x$, where we denote the input prices by p in $\mathbb{R}^I_{++}$.

[14] A quick way of verifying this is to assume that g is sufficiently smooth and show that $\partial^2\phi/\partial a \partial k \leq 0$, but it is not difficult to give a discrete proof that dispenses with differentiability.

[15] We are grateful to Eddie Dekel, whose queries inspired us to look at this issue more closely.

The benefit $B(q)$ may be interpreted as the revenue derived from selling q units of the good, which is generally non-linear in q if the firm has monopoly power. Alternatively, the firm could face a risky output price s, in which case

$$B(q) := \int_0^\infty u(sq)d\mu(s),$$

where $u : \mathbb{R}_+ \to \mathbb{R}$ is the Bernoulli index that summarises the firm's attitude towards risk, while μ is the probability distribution over the output price s.

We know that the factors are complements if the map from x to $B\big(f(x)\big)$ is supermodular. However, unless we make further assumptions about B, we cannot obtain such a conclusion even if f is supermodular. Nonetheless, with suitable assumptions on f alone, we can guarantee the the firm will *raise its output as the price of a factor falls.*

Given the production function f, the firm's cost function $C : \mathbb{R}^I_{++} \times \mathbb{R}_+ \to \mathbb{R}$ is

$$C(p, q) := \min\big\{p \cdot y : f(y) \geq q\big\}. \tag{6}$$

To keep the exposition short, suppose that, for any $q \geq 0$, there is some $x \in \mathbb{R}^I_+$ such that $f(x) = q$. Moreover, let C be well-defined, for all $p \in \mathbb{R}^I_{++}$ and $q \geq 0$. Hence, the firm's optimisation is equivalent to choosing an output $q \geq 0$ that maximises $B(q) - C(p, q)$.

We wish to find conditions on function f under which the firm's output increases when the p_i price of factor i falls. By the Topkis MCS Theorem, it suffices for the cost function C to be supermodular in (p_i, q). It is clear from the argument in Quah (2007) that the crucial condition to imply this property on C is supermodularity and i-concavity of the production function f.[16] However, the argument in that paper is rather roundabout — it uses the Envelope Theorem and relies on the differentiability of C, as well as various (rather strong) ancillary assumptions. We now provide a direct proof of this result.

Proposition 3. *Suppose the function f is continuous, strictly increasing, supermodular, and i-concave for some factor $i \in I$. Then, function C, defined in (6), is supermodular in (p_i, q), for any prices p_{-i} of the remaining factors.*[17]

Proof. Given our results, there is a natural proof strategy for this claim. Suppose that the correspondence $\Gamma_i : \mathbb{R}_{++} \times \mathbb{R} \to \mathbb{R}^I_+$, given by

$$\Gamma_i(p_i, q) := \big\{(p_i y_i, y_{-i}) : f(y) \geq q\big\},$$

[16] A function f is i-concave if, for any fixed x_i, it is concave function of x_{-i}.

[17] If we wish to guarantee that the firm raises its output whenever the price of any factor falls, then we require f to be i-concave for all $i = 1, \dots, I$. Note that it is possible for a function to be i-concave for all i without being concave. For example, $f(x_1, x_2) = x_1 x_2$ is i-concave for $i = 1, 2$, but it is not concave.

is lower supermodular. Then, by the Main Theorem, function $v : \mathbb{R}_{++} \times \mathbb{R}_{+} \to \mathbb{R}$,

$$v(p_i, q) := \min \big\{ (1, p_{-i}) \cdot (z_i, z_{-i}) : z \in \Gamma_i(p_i, q) \big\},$$

is supermodular. Since $v_i(p_i, q) = C\big((p_i, p_{-i}), q\big)$, C is supermodular in (p_i, q). We show in the Appendix that the conditions on f guarantee that Γ_i is lower supermodular. $\square$

Application 3: Supermodular value theorems

Let X and T be lattices and Z a sublattice of $X \times T$. As in Example 5, we denote by X_Z the set of elements in X for which the sectional set $Z(x) := \{t \in T : (x,t) \in Z\}$ is non-empty. Topkis (1978) has shown that if $f : Z \to \mathbb{R}$ is a supermodular function then the envelope function, $v : X_Z \to Y$, given by

$$v(x) = \max \big\{ f(x,t) : t \in Z(x) \big\}$$

is in turn a supermodular function; we shall refer to this result as the *Topkis Value Theorem*. It has important applications, for example in the theory of dynamic optimization (see Hopenhayn and Prescott, 1992), which we shall discuss in the next section. This result can be obtained as a straightforward application of the Main Theorem since we have shown in Example 5 that the correspondence Γ given by $\Gamma(x) = \big\{ f(x,t) : t \in Z(x) \big\}$ is upper supermodular. It is also clear from this perspective that *any* combination of properties that leads to the upper supermodularity of Γ will guarantee the supermodularity of v. We now present an alternative set of conditions that is useful in applications.

We specialize to the case where X is a convex sublattice of $\mathbb{R}^n$, $T \subseteq \mathbb{R}$, and $Z = X \times T$, so that $Z(x) = T$, for all $x \in X$. The function $f : X \times T \to \mathbb{R}$ is required to be supermodular in X, but it is not required to be jointly supermodular in (x,t). Instead we assume that f has *single-crossing differences in* (t,x). It is straightforward to check that if f is jointly supermodular in (x,t) then it is supermodular in x and has single crossing differences in (t,x), however, as we shall see in Example D discussed below, the converse is not true. Nonetheless, we know from the Milgrom-Shannon MCS Theorem, that this property is sufficient to guarantee that $\arg\max_{t \in T} f(x,t)$ is increasing in x (in the strong set order), and it seems intuitive that a property which is useful in extending the Topkis MCS Theorem should also be useful in extending the Topkis Value Theorem.

Proposition 4. *Let X be a convex sublattice of $\mathbb{R}^I$ and $T \subseteq \mathbb{R}$. Moreover, suppose that the function $f : X \times T \to \mathbb{R}$ is a continuous and supermodular function of x, for all $t \in T$, and that it obeys single crossing differences in (t, x). Then correspondence $\Gamma : X \to \mathbb{R}$, given by $\Gamma(x) = \{f(x,t) : t \in T\}$, is upper supermodular.*[18] *Consequently, function $v(x) = \max\{f(x,t) : t \in T\}$ is supermodular in x.*

Example D. Suppose a firm has access to production technologies drawn from a set $\{f(\cdot,t) : t \in T\}$ for some function $f : X \times T \to \mathbb{R}$, where X is a convex sublattice of $\mathbb{R}^n$ and $T \subset \mathbb{R}$. Each technology $f(\cdot,t)$ maps an input vector x to $f(x,t)$, where the latter can be interpreted either as the level of physical output or as revenue. Assuming that the firm can switch costlessly among the technologies, its production function is given by

$$F(x) := \max\{f(x,t) : t \in T\}. \tag{7}$$

Proposition 4 tells us that F is a supermodular function if $f(x,t)$ is a continuous and supermodular function of x and has single crossing differences in (t, x). Assuming these conditions on f and that the firm chooses x to maximize profit $F(x) - p \cdot x$, where p is the vector of input prices, we know that a fall in the price of factor i leads to an increase in the optimal x. Furthermore, since the technology type t is chosen optimally given x and thus will be increasing with x, by the Milgrom-Shannon MCS Theorem, we conclude that the *optimal t increases with a fall in the price of i*.

To be specific, consider the case where $f((k,\ell),t) := A(t)k^{\alpha(t)}\ell^{\beta(t)}$, where $A(t)$ takes positive values and $\alpha(t)$ and $\beta(t)$ are strictly positive and increasing in t. Therefore, the family of functions $\{f(\cdot,t) : t \in T\}$ consist of Cobb-Douglas production functions (with capital and labour as inputs) that are weakly ordered with respect to the output elasticities. Each function $f(\cdot,t)$ is supermodular and, for any $t' \geq t$, the difference

$$\Delta(k,\ell) := A(t')k^{\alpha(t')}\ell^{\beta(t')} - A(t)k^{\alpha(t)}\ell^{\beta(t)}$$

is single crossing. Indeed, for any $k > 0$, it is straightforward to show that $\Delta(k,\ell') \geq (>)\, 0$ implies $\Delta(k,\ell'') \geq (>)\, 0$ for any $\ell'' > \ell' > 0$ and the analogous property holds if ℓ is held fixed. We conclude, by Proposition 4 that function $F : \mathbb{R}^2_{++} \to \mathbb{R}_+$ is supermodular.

However, $f((k,\ell),t)$ is *not* jointly supermodular in $((k,\ell),t)$ and thus the Topkis Value Theorem is not applicable. To see this, fix $k > 0$ and notice that $\Delta(k,\cdot)$ takes the

[18] In fact the conclusion still holds if we replace single crossing differences with the weaker property that, for all $t' \geq t$ and $x' \geq x$, if $f(x,t') \geq f(x,t)$ then $f(x',t') \geq f(x',t)$.

value zero at $\ell = 0$ and (at most one) other point $\ell^* > 0$, with $\Delta(k,\ell) < 0$ for $0 < \ell < \ell^*$ and $\Delta(k,\ell) > 0$ for $\ell < \ell^*$. Hence, $\Delta(k,\ell)$ *cannot* be an increasing function of ℓ, which means that $f\big((k,\ell),t\big)$ is *not* jointly supermodular in $\big((k,\ell),t\big)$.

Even though Proposition 4 weakens the properties on f relative to Topkis' result, it is not a generalization of the latter. In particular, while it assumes that $Z = X \times T$, Topkis' result allows Z to be a sublattice of $X \times T$, and there are counter examples to show that the proposition *cannot* be extended to an arbitrary sublattice of $X \times T$.[19] However, it is possible to extend Proposition 4 to the case where T is a product set (and hence a sublattice) in a Euclidean space. For such an extension, in addition to f having single crossing differences in (t,x), we also require f to be a quasi-supermodular function of t. Note that, by the Milgrom-Shannon MCS Theorem, these are precisely the properties which guarantee that $\arg\max_{t\in T} f(x,t)$ is increasing in x.

Proposition 5. *Let X be a convex sublattice of $\mathbb{R}^n$ and $T = \times_{i=1}^m T_i$, where T_i are compact subsets of $\mathbb{R}$. Suppose that the function $f : X \times T \to \mathbb{R}$ has the following properties:* (i) *it is continuous in (x,t);* (ii) *supermodular in x, for all $t \in T$;* (iii) *quasisupermodular in t, for all $x \in X$; and* (iv) *obeys single crossing differences in (t,x). Then, function $v(x) = \max\big\{f(x,t) : t \in T\big\}$ is a continuous and supermodular on X.*

The proof of Proposition 5 involves the repeated application of Proposition 4. It also makes use of Proposition 6 below, a result which may be of interest in itself; Proposition 6 could be thought of as the exact quasisupermodular analog to the Topkis Value Theorem since it states that the value function preserves the quasisupermodularity of the objective function. On the other hand, Propositions 4 and 5 are hybrid results where one obtains a supermodular value function, while relaxing supermodularity of the objective function.

Proposition 6. *Let X and T be lattices and Z be a sublattice of $X \times T$, with sets X_Z and $Z(x)$ defined as in Example 5. Moreover, suppose that function $f : Z \to \mathbb{R}$ is quasisupermodular and admits a non-empty $\arg\max_{t\in Z(x)} f(x,t)$, for all $x \in X_Z$. Then, function $v(x) = \max\big\{f(x,t) : t \in Z(x)\big\}$ is quasisupermodular on X_Z.*

[19] Let $X = \{x, x', x \wedge x', x \vee x'\}$ and $T = \{t, t'\}$, where x and x' are two unordered vectors in $\mathbb{R}^2$, and t and t' are elements of $\mathbb{R}$, with $t' > t$. Let $Z = \big\{(x,t),(x',t'),(x\wedge x',t),(x \vee x',t')\big\}$ and define function $f : X \times T \to \mathbb{R}$ as follows: $f(x,t') = 2$, for all $x \in X$, $f(x,t) = f(x',t) = 1$, $f(x \vee x',t) = 1.5$, and $f(x \wedge x',t) = 0.5$. Hence, function f is a supermodular in x and obeys single crossing differences in (t,x), but it is *not* jointly supermodular in (x,t), since $f(x',t') + f(x,t) > f(x \wedge x',t) + f(x \vee x',t')$. In particular, as $X_Z = X$, we have $v(x) + v(x') = f(x,t) + f(x',t')$ and $v(x \wedge x') + v(x \vee x') = f(x \wedge x',t) + f(x \vee x',t')$, which contradicts that v is a supermodular function.

Application 4: Shifts in multi-prior beliefs

Consider an agent who chooses an action $x \in X$ before the realisation of some state $s \in S$, where X and S are both subsets of $\mathbb{R}$. Given x, the agent's utility is $f(x, s)$ whenever state s is realised. Assuming that λ is a probability distribution over S, the agent chooses x to maximise the expected utility $\int_S f(x, \tilde{s}) d\lambda(\tilde{s})$. If f is a supermodular function and the agent is allowed to choose her action *after* observing the state, then we know that her action will increase with the state. Therefore, it is intuitive that under the same condition on f, if the agent has to make a decision *before* the state is realised, then she will pick a higher action if she thinks that higher states are more likely. This turns out to be true; more precisely, it can be shown that a first order stochastic dominance shift in the distribution of s will indeed lead to a higher optimal action.

In this application we extend this result to the case where the agent has non-expected utility preferences. The *maxmin* model of Gilboa and Schmeidler (1989) allows the agent to be *ambiguity averse*; in this case, the agent behaves as though her belief over S consists of a *set* Λ of probability distributions over S, with the agent's preference having the representation $\min_{\lambda \in \Lambda} \int_S f(x, \tilde{s}) d\lambda(\tilde{s})$. The α-maxmin model (see Ghirardato, Maccheroni, and Marinacci, 2004), which generalizes the maxmin model, allows for both ambiguity averse and ambiguity loving behaviour, with the agent's utility having the form

$$\alpha \min_{\lambda \in \Lambda} \int_S f(x, s) d\lambda(s) + (1 - \alpha) \max_{\lambda \in \Lambda} \int_S f(x, s) d\lambda(s)$$

for some $\alpha \in [0, 1]$. In this application, we show that an agent with α-maxmin preferences will take a higher action when there is a shift in her beliefs towards higher states.

To be precise, let $\triangle_S$ denote the set of probability distributions over S. Since $S \subseteq \mathbb{R}$, the set $\triangle_S$ is a lattice if we rank elements in $\triangle_S$ by first order stochastic dominance. Indeed, for any measures λ, λ' in $\triangle_S$, we have $(\lambda \wedge \lambda')(s) = \max\left\{\lambda(s), \lambda'(s)\right\}$ and $(\lambda \vee \lambda')(s) = \min\left\{\lambda(s), \lambda'(s)\right\}$, for all $s \in S$. Therefore, subsets of $\triangle_S$ could in turn be compared using the strong set order induced by first order stochastic dominance. This way of comparing sets of distributions captures precisely the notion of greater belief in higher states which is needed for our purposes.

In the next result, we introduce a parameter such that higher states are considered more likely when the parameter takes a higher value. Formally, the set of parameters is $T \subseteq \mathbb{R}$ and beliefs are represented by a correspondence $\Lambda : T \to \triangle_S$ which is increasing in the strong set order induced by first order stochastic dominance, i.e., set $\Lambda(t')$ dominates $\Lambda(t)$ in the strong set order, for any $t' \geq t$.As a simple example, suppose that λ_m, $\lambda_M : T \to \triangle_S$ are two increasing functions, with $\lambda_M(t)$ first order stochastically dominating $\lambda_m(t)$ for all $t \in T$. Then the set $\Lambda(t)$ of all distributions lying between $\lambda_m(t)$ and $\lambda_M(t)$ is a sublattice of $\triangle_S$ and the correspondence Λ is increasing in the strong set order.

Proposition 7. *Let X, S and T be subsets of $\mathbb{R}$, with S being finite. Whenever function $f : X \times S \to \mathbb{R}$ is supermodular and correspondence $\Lambda : T \to \triangle_S$ is increasing with respect to the strong set order induced by first order stochastic dominance, then functions*

$$v(x,t) := \min \left\{ \int_S f(x,s) d\lambda(s) : \lambda \in \Lambda(t) \right\} \tag{8}$$

$$w(x,t) := \max \left\{ \int_S f(x,s) d\lambda(s) : \lambda \in \Lambda(t) \right\}. \tag{9}$$

are supermodular in (x,t). Consequently, its convex combination $\alpha v(x,t) + (1-\alpha) w(x,t)$, is supermodular function, for all $\alpha \in [0,1]$.

Remark. We show in the Appendix that the conclusions remain true if S is a compact interval of $\mathbb{R}$ and function $f(x,\cdot)$ is Riemann–Stieltjes integrable over S with respect to each $\lambda \in \Lambda(t)$, for all $x \in X$ and $t \in T$. In particular, this holds if any of the following conditions is satisfied: (a) $f(x,s)$ is continuous in $s \in S$; (b) $f(x,s)$ is bounded on S and has only finitely many discontinuities in s and all distributions in $\Lambda(t)$ are atomless; or (c) $f(x,s)$ is bounded on S and monotone and all distributions in $\Lambda(t)$ are atomless.

Proof. The idea of the proof is to show that the objective function is equivalent to the infimum (or supremum) of some positive linear functional over a submodular correspondence. The conclusion then follows from the Main Theorem (∗). First, denote $S = \{s_i\}_{i=1}^{\ell+1}$ such that $s_1 < s_2 < \ldots < s_{\ell+1}$. For any $x \in X$ and $\lambda \in \Lambda(t)$,

$$\begin{aligned} \int_S f(x,s) d\lambda(s) &= f(x,s_1)\lambda(s_1) + \sum_{i=1}^{\ell} f(x,s_{i+1}) \big[\lambda(s_{i+1}) - \lambda(s_i)\big] \\ &= f(x,s_{\ell+1})\lambda(s_{\ell+1}) + \sum_{i=1}^{\ell} \big[f(x,s_i) - f(x,s_{i+1})\big] \lambda(s_i) \\ &= f(x,s_{\ell+1}) - \sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i), \end{aligned} \tag{10}$$

since $\lambda(s_{\ell+1}) = 1$, while $\delta_i(x) := \big[f(x, s_{i+1}) - f(x, s_i)\big]$, for all $i = 1, \ldots, \ell$. This observation allows us to represent function v by

$$v(x,t) = f(x, s_{\ell+1}) - \max\left\{\sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i) : \lambda \in \Lambda(t)\right\},$$

which is supermodular whenever $\bar{v}(x,t) := \max\big\{\sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i) : \lambda \in \Lambda(t)\big\}$ is submodular. By the Main Theorem $(*)$, this holds once the correspondence $\Gamma : X \times T \to \mathbb{R}^\ell$,

$$\Gamma(x,t) := \Big\{y \in \mathbb{R}^\ell : y_i = \delta_i(x)\lambda(s_i) \text{ for some } \lambda \in \Lambda(t)\Big\},$$

is lower submodular (with respect to the coordinate-wise partial order on $\mathbb{R}^\ell$). Analogously, we show that w is supermodular if $\bar{w}(x,t) := \min\big\{\sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i) : \lambda \in \Lambda(t)\big\}$ is a submodular function. By the Main Theorem $(*)$, this holds if Γ is an upper submodular correspondence. See the Appendix. □

Example E. 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 $S \subset \mathbb{R}_+$. The investor's beliefs over the return of the risky asset is captured by the correspondence $\Lambda : T \to \triangle_S$, where $\triangle_S$ is the space of probability distributions over S. To capture the idea that a higher t corresponds to greater optimism, we assume that Λ is increasing in t with respect to the strong set order induced by first order stochastic dominance.

The investor chooses to invest $x \in X \subset \mathbb{R}$ 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 always solvent, i.e., $xs + (m - x)r \geq 0$ 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

$$v(x,t) := \min\left\{\int_S u\big(xs + (m-x)r\big)d\lambda(s) : \lambda \in \Lambda(t)\right\}. \tag{11}$$

By Proposition 7, v is supermodular if $f(x,s) := u\big(xs + (m-x)r\big)$ is supermodular. Assuming that u is strictly increasing, concave, and twice continuously differentiable, it is straightforward to check that f is supermodular if the coefficient of relative risk aversion of u is less than 1. Therefore, with this condition on u, we can apply the Topkis 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]

[20] We are not the first to discuss comparative statics of the portfolio choice model under ambiguity. For

The next example has a somewhat different flavour from Example E: it has both x and t as choice variables and exploits the fact that Proposition 7 guarantees that the maxmin utility in (8) is a supermodular function, which means that its sum with another supermodular function will again be supermodular.

Example F. 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 is given by a multi-prior set $\Lambda(t)$ of probability distributions over the set $S \subseteq \mathbb{R}_+$. We assume that higher advertising leads to greater optimism in the sense that the correspondence $\Lambda : T \to \triangle_S$ is increasing in the strong set order. The price of the good is assumed to be fixed at 1. 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; therefore the profit in period 2 is $\pi(x, s)$, where $\pi(x, s) = s$ if $s \le x$ (so undemanded goods can be freely disposed) and $\pi(x, s) = s - \kappa(s - x)$ if $s > x$; $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ should be interpreted as the cost incurred for producing the additional units to meet demand in period 2. Notice that $\pi(x, s)$ need not be increasing in s. The firm chooses $x \ge 0$ and $a \ge 0$ in period 1 to maximize

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

It is straightforward to check that the function π is supermodular if κ is increasing and convex, with $\kappa(0) = 0$.[21] Given this, Proposition 7 guarantees that

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

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 Topkis 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.

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.

[21] For $x' \ge x$, the difference term $\pi(x', s) - \pi(x, s)$ equals 0 if $s \le x$, $\kappa(s - x)$ if $x < s \le x'$, and $\kappa(s - x) - \kappa(s - x')$ if $s > x'$. Under the assumptions on κ, this is increasing in s.

Application 5: Variational and multiplier preferences

It is possible to extend Proposition 7 to cover a broader class of preferences. Maccheroni, Marinacci, and Rustichini (2006) introduce and axiomatise a generalisation of the Gilboa-Schmeidler maxmin model, called *variational preferences*. In this class of preferences, the utility of some action x is $v(x) = \min_{\lambda \in \triangle_S} \left\{ \int_S f(x,s) d\lambda(s) + c(\lambda) \right\}$. Very loosely speaking, the agent's utility from the action x is obtained by minimizing over a set of probability distributions. Unlike the maxmin model where the agent is restricted to a subset of Δ_S, any distribution in Δ_S could be 'picked' in the variational preferences model, though each distribution λ is associated with a different cost $c(\lambda)$. For a detailed discussion of these preferences see Maccheroni, Marinacci, and Rustichini (2006) or the survey of Epstein and Schneider (2010).

In the following result, the cost functions are parametrised by $t \in T \subseteq \mathbb{R}$ and we identify conditions under which the agent's utility is supermodular in (x, t).

Proposition 8. *Let X, S and T be subsets of $\mathbb{R}$ and assume that S is a finite set. If the function $f : X \times S \to \mathbb{R}$ is supermodular and the function $c : \triangle_S \times T \to \mathbb{R}_+ \cup \{\infty\}$ is submodular.*[22] *then the function $v : X \times T \to \mathbb{R}$ defined by*

$$v(x,t) := \min \left\{ \int_S f(x,s) d\lambda(s) + c(\lambda, t) : \lambda \in \triangle_S \right\} \tag{12}$$

is a supermodular function.[23]

This result is quite natural. When c is submodular, the marginal cost of choosing a higher λ, with respect to first order stochastic dominance, falls as t increases. This guarantees that $\Lambda_*(x,t)$, the set of distributions that solves the minimization problem in (12), increases with t in the strong set order.[24] In other words, when evaluating the ex ante utility of a given action x, a higher distribution is used when t is higher. When f is supermodular, higher actions are favoured at higher states, so it is intuitive that the ex ante utility v will favour higher actions when t is higher.

[22] This is with respect to the product order of the first order stochastic dominance order on $\triangle_S$ and the usual order on $T \subset \mathbb{R}$.

[23] This claim remains true when S is a compact interval and, for each (x,t), $f(x,\cdot)$ is Riemann-Stieltjes integrable with respect to all $\lambda \in \Lambda(t)$. See the Remark following Proposition 7.

[24] From the proof of Proposition 8, we obtain $\Lambda_*(x,t) = \arg\max_{\lambda \in \Lambda(t)} \left\{ \sum_{i=1}^{\ell} \delta_i(x) \lambda(s_i) - c(\lambda, t) \right\}$. See equation (A5) in the Appendix. It is easy to check that the objective function is supermodular in (λ, t) when c is submodular and so, by the Topkis MCS Theorem, $\Lambda_*(x,t)$ increases with t.

One part of Proposition 7 can be thought of as a special case of Proposition 8. Indeed, given a correspondence $\Lambda : T \to \triangle_S$, if we define the function c by

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

then $v(x, t)$, defined in (12), takes the maxmin form given in (8). Furthermore, if Λ is increasing in the sense of the strong set order, then one could check that c is submodular and thus the supermodularity of v is guaranteed by Proposition 8.

Another prominent example of this class of models are *multiplier preferences*, introduced in Sargent and Hansen (2001) and axiomatised by Strzalecki (2011a). In this case, the cost c is given by $c(\lambda, t) := \theta R\big(\lambda \| \lambda^*(\cdot, t)\big)$, for some $\theta \geq 0$ and $\lambda^*(\cdot, t) \in \triangle_S$, where

$$R\big(\lambda \| \lambda^*(\cdot, t)\big) := \sum_i \pi_i \ln\left(\frac{\pi_i}{\pi_i^*(t)}\right)$$

is the *relative entropy*.[25] Note that π_i ($\pi_i^*(t)$) refers to the probability of state i in the distribution λ (respectively, $\lambda^*(\cdot, t)$). 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 (12). The minimisation over all possible distributions reflects aversion to misspecification of the model or ambiguity.

In the Appendix, we prove the following result.

Proposition 9. *Let S and T be subsets of $\mathbb{R}$ and assume that S is a finite set. For any fixed $\lambda^*(\cdot, t)$, the relative entropy $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is a submodular function of $\lambda \in \triangle_S$. Furthermore, it is a submodular function of (λ, t) if $\lambda^*(\cdot, t)$ is increasing in t with respect to the monotone likelihood ratio order.*

Remark. The property on $\lambda^*(\cdot, t)$ means the following: if $t' \geq t$, then the ratio $\pi_i^*(t')/\pi_i^*(t)$ is increasing with i. It is straightforward to check that this condition implies that $\lambda(\cdot, t')$ dominates $\lambda(\cdot, t)$ with respect to first order stochastic dominance.

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

By combining Propositions 8 and 9, we conclude that $v(x,t)$ is supermodular if function f is supermodular and the agent has multiplier preferences, with the benchmark distribution increasing with t in the sense of the monotone likelihood ratio order. In other words, the marginal utility of a higher action becomes greater when the benchmark distribution shifts in favour of higher states. In Examples E and F we assume that the agent has maxmin preferences; it is clear that, by appealing to Proposition 8 (instead of Proposition 7), the conclusions in those examples will continue to hold, mutatis mutandi, if the agent has variational or, more specifically, multiplier preferences.

5 Monotone methods in dynamic programming

In an influential paper, Hopenhayn and Prescott (1992) used the tools of monotone comparative statics to analyse stationary dynamic optimization problems. In this section, we show how the results we have developed could also be fruitfully in that context.

We consider an agent who faces a stochastic control problem where X and S are the sets of endogenous and exogenous state variables. To keep the exposition simple, we shall assume that X is 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; $x_{\tau+1}$ is chosen from a non-empty feasible set which may depend on the current state and which we denote by $\Phi(x_\tau, s_\tau) \subseteq X$. The single-period return is given by the function $u : X \times S \times X \to \mathbb{R}$; $u(x,s,y)$ 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 payoffs are discounted 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 u and the continuity of Φ (see Theorem 9.6 in Stokey, Lucas, and Prescott, 1989 for details) — this problem admits a recursive representation and $v = v^*$ is the unique solution to the Bellman equation

$$v(x,s) = \max\left\{u(x,s,y) + \beta\int_S v(y,\tilde{s})d\lambda(\tilde{s},s) : y \in \Phi(x,s)\right\}, \tag{13}$$

where $\lambda(\cdot, s)$ gives the distribution on S in the following period, conditional on s. The function v is bounded and continuous and, if we define the operator $\mathscr{T}$ by

$$\big(\mathscr{T}w\big)(x,s) = \max\left\{u(x,s,y) + \beta\int_S w(y,\tilde{s})d\lambda(\tilde{s},s) : y \in \Phi(x,s)\right\}, \tag{14}$$

that maps the space of bounded and continuous real-valued functions on $X \times S$ into itself, then beginning at *any* bounded and continuous function $w : X \times S \to \mathbb{R}$, $\mathscr{T}^n w$ converges uniformly to v^*. Furthermore, for all $(x,s) \in X \times S$, the set

$$\gamma(x,s) := \arg\max\left\{u(x,s,y) + \beta\int_S v^*(y,\tilde{s})d\lambda(\tilde{s},s) : y \in \Phi(x,s)\right\} \tag{15}$$

is non-empty and compact and the correspondence γ is upper hemi-continuous. We shall refer to any optimal control problem in which v^* and γ have the properties we have just listed in this paragraph as a *well-behaved* problem.

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

$$\alpha^*(x,s,y) := u(x,s,y) + \beta\int_S v^*(y,\tilde{s})d\lambda(\tilde{s},s) \tag{16}$$

is supermodular in y and has increasing differences in $\big(y,(x,s)\big)$. By the Topkis MCS Theorem, $\gamma(x,s)$ is a compact sublattice of X and it is increasing in (x,s).[26] This in turn guarantees the existence of an increasing policy function; for example, one could choose $g(x,s) = \max\gamma(x,s)$.[27] HP also show that g is Borel measurable. Lastly, the policy function g induces a Markov process on $X \times S$, where, for measurable sets $A \subseteq X$ and

[26] Condition (ii) on Φ guarantees that $\Phi(x,s)$ is sublattice of X and that it increases with (x,s) in the strong set order. Given with the properties on α^*, we know that $\gamma(x,s)$ is a sublattice and that it increases with (x,s); this follows from a more general version of the Topkis MCS Theorem (than the one stated in Section 4) that allows for increasing constraint sets. See Topkis (1978).

[27] In other words, $g(x,s)$ is the largest element of $\gamma(x,s)$ with respect to the product order; g is well-defined because γ is compact-valued and a sublattice.

$B \subseteq S$, the probability of $A \times B$ conditional on (x, s) is the probability of B conditional on s if $g(x, s) \in A$, and it is zero otherwise. HP make use of the monotonicity of g to guarantee that this Markov process has a stationary distribution.[28]

We now apply our results to build two models with monotone policy functions.

Application 6: Conditional increasing differences

While there are many applications of the Milgrom-Shannon MCS Theorem, its use in dynamic programming problems has been limited because in these problems conditions such as supermodularity and increasing differences are typically imposed on primitives. There are two related reasons for this. First, the Topkis Value Theorem plays a crucial role in this class of applications, and there are hitherto no versions of that theorem where the supermodularity and/or increasing differences assumptions are relaxed. Second, in choosing her action, the agent takes into account both the instantaneous payoff *and* the continuation payoff; even when single crossing differences is imposed on the instantaneous payoff, it does not generally lead to that property holding on the agent's overall objective function, because single crossing differences (unlike increasing differences) is not preserved by aggregation. In this application, we identify a property lying between increasing and single crossing differences which *does* have some aggregative features. For payoff functions with this property, we could guarantee that policy functions are monotone; the proof uses our extension of the Topkis Value Theorem because the weaker assumptions render the original theorem inapplicable.

Function u obeys *conditional increasing differences* if for any $y' \geq y$ and (x, s) such that $u(x, s, y') - u(x, s, y) \geq 0$ implies $u(x', s', y') - u(x', s', y) \geq u(x, s, y') - u(x, s, y)$. This property, which seems at first blush rather special, in fact, occurs quite naturally. For example, it is sometimes more natural in applications to assume that u has log increasing differences in $\big(y, (x, s)\big)$, i.e., the ratio $u(x, s, y')/u(x, s, y)$ is increasing in (x, s) whenever $y' \geq y$. If, in addition, we assume that u is increasing in (x, s), then u obeys conditional increasing differences, even though it may not obey increasing differences.

[28] 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, more recently, Stachurski and Kamihigashi (2014) also discuss uniqueness and other issues relating to the stationary distribution.

Proposition 10. *Suppose that* $X = \times_{i=1}^{\ell} X_i$ *where* X_i *are compact intervals of* $\mathbb{R}$*,* S *is a convex and compact subset of some Euclidean space,* $\Phi(x,s) = X$ *for all* (x,s) *in* $X \times S$*, and the optimal control problem is well-behaved. We assume that* $u(x,s,y)$ *is jointly continuous in* (x,s,y)*, decreasing in* y*, supermodular in* x*, supermodular in* y*, and has conditional increasing differences in* $\big(y,(x,s)\big)$*. Then, the value function* $v^*(x,s)$ *is supermodular in* x *and has increasing differences in* (x,s)*, while correspondence* γ*, defined in (15), is sublattice-valued and increasing in the strong set order, and* $g : X \times S \to \mathbb{R}$*, given by* $g(x,s) := \max \gamma(x,s)$*, is an increasing, Borel measurable policy function.*

Remark. If X and S are both compact intervals of $\mathbb{R}$, the conditions on u simplify to requiring u to be jointly continuous in (x,s,y), decreasing in y, supermodular in (x,s) and to have conditional increasing differences in $\big(y,(x,s)\big)$.

This proposition does not provide a true generalization of HP since we impose conditions which are in some respects stronger; in particular, we require more structure on X and S and also assume that the constraint set at each period does not vary with (x,s). However, unlike that approach, we do not require $u(x,s,y)$ to be jointly supermodular in (x,y) and to have increasing differences in $\big((x,y),s\big)$; these conditions imply that $u(x,s,y)$ is supermodular in x, supermodular in y, has increasing differences in (x,s) and has conditional increasing differences in $\big(y,(x,s)\big)$, but the converse is not true.

The proof of Proposition 10 is in the Appendix. It suffices to highlight its notable features. Given the weaker conditions imposed on u, we cannot guarantee that $v^*(x,s)$ is supermodular in x and has increasing differences in (x,s) by applying the Topkis Value Theorem. However, these properties can still be ensured by applying Proposition 5. The objective function $\alpha^*(x,s,y)$, see (16), while it is supermodular in y, only obeys single crossing differences in $\big(y,(x,s)\big)$ rather than increasing differences. Nonetheless, by the Milgrom-Shannon MCS Theorem, this is sufficient to guarantee the monotonicity of γ.

Example G. We study the problem of a monopolist selling a single good over time in an uncertain environment. We assume that the firm's production process benefits from learning by doing, so that current output lowers the marginal cost of production in the next period. This dependence introduces a dynamic element to the firm's decision problem: in deciding how much to produce this period, the firm must also consider its impact on cost in the following period and the extent to which it could respond to market conditions prevailing then.

In each period, the firm faces an inverse demand function $P(y, s)$, where y denotes the quantity sold and the parameter s summarises the current state of the market. The production cost $c(x, y)$ depends on the current output level y as well as on the output x in the preceding period. In each period, the firm observes s, which is drawn from a compact interval S, and chooses output y from a compact interval $X \subset \mathbb{R}_+$. The profit of the firm in that period is given by the function $u : X \times S \times X \to \mathbb{R}$, where

$$u(x, s, y) := yP(y, s) - c(x, y).$$

We assume that P and c are both continuous functions, which guarantees that u is also continuous. To apply Proposition 10, we need to guarantee that u is decreasing in x, supermodular in (x, s) and has conditional increasing differences in $\big(y, (x, s)\big)$. It is clearly supermodular in (x, s) since it is additively separable in those variables. To guarantee the other properties we require further assumptions on P and c.

We assume that $c(x, y)$ *is increasing in* (x, y) *and submodular in* (x, y). That current costs should increase with past production x is quite natural if we interpret these added costs as payments (to certain suppliers, etc.) which only need to be made one period after production; we could also think of them as costs associated with providing service support or guarantees for products sold in the previous period. Clearly, this assumption guarantees that u is decreasing in x. The submodularity of c means that the marginal cost of current production decreases with the output in the previous period, which models our assumption that there is learning by doing. A simple example of a function that obeys both properties is $c(x, y) = \psi(kx + y)$, where $k > 0$ and ψ is increasing and concave.

We also assume that $P(y, s)$ *is increasing in* s *and log-supermodular*. The latter is an easily interpretable property because it is equivalent to saying that the own-price elasticity of demand for this product falls as s increases, i.e., the reciprocal of

$$-\frac{y}{P(y, s)} \frac{\partial P}{\partial y}(y, s)$$

decreases with s. In short, a higher s benefits the firm by raising the clearing price at any output level and by making the firm's demand less sensitive to price.

These assumptions guarantee that $u(x, s, y)$ obeys conditional increasing differences in $\big(y, (x, s)\big)$. Suppose that for $y' \geq y$ and (x, s) we have $u(x, s, y') - u(x, s, y) \geq (>)\ 0$.

By monotonicity of c, this implies that $y'P(y',s) \geq (>) yP(y,s)$. Note that

$$\begin{aligned} & y'P(y',s') - yP(y,s') \\ & \quad = yP(y,s')\left(\frac{y'P(y',s')}{yP(y,s')} - 1\right) \geq yP(y,s)\left(\frac{y'P(y',s')}{yP(y,s')} - 1\right) \\ & \qquad \geq yP(y,s)\left(\frac{y'P(y',s)}{yP(y,s)} - 1\right) = y'P(y',s) - yP(y,s). \end{aligned}$$

The first inequality holds because $P(y,s') \geq P(y,s)$ and the term in the large brackets is positive while the second inequality follows from the log-supermodularity of P.[29] By submodularity of c, we have $-\big[c(x',y') - c(x',y)\big] \geq -\big[c(x,y') - c(x,y)\big]$. Thus, we obtain

$$u(x',s',y') - u(x',s',y) \geq u(x,s,y') - u(x,s,y).$$

With these assumptions on P and c in place, and assuming that the transition function $\lambda(\cdot, s)$ is increasing in s with respect to first order stochastic dominance, Proposition 10 guarantees that the firm's policy function is increasing in (x,s). The intuition for this result is as follows: when (x,s) increases, the current-period payoff of choosing a higher output increases because demand is less elastic and also because the marginal cost of output is lower; furthermore, while raising output lowers the firm's continuation value, the drop is smaller when the exogenous state is higher, which also encourages raising output. Lastly, the monotonicity of the firm's policy function imply that there is a stationary joint distribution on (x,s).

Application 7: Dynamic maxmin preferences

We consider a stochastic control problem identical to the one described at the beginning of this section. Since at each period τ, the exogenous variable takes a value drawn from the set S, the set of all possible realizations of the exogenous variable over time is given by S^∞. If the agent is an expected utility-maximizer, then she behaves as though she is guided by a distribution over S^∞; to obtain the utility of a given plan of action, the agent first evaluates the discounted utility on every possible path, i.e., over every element in S^∞, and then takes the average across paths, weighing each path with its probability.

[29] The argument here is similar to, and inspired by, the argument made in Amir (1996) to guarantee that, in a static Cournot game, a firm's best responds to higher output from other firms by reducing its output. Amir (1996) shows that it suffices for the firm's cost function to be increasing in output and for the inverse demand function for the homogeneous product sold in the market to be log-concave.

On the other hand, when the agent has a maxmin preference, her behaviour can be modelled by a *set* of distributions $\mathcal{B}$ over S^∞. The utility of a plan is then given by the minimum of the expected discounted utility for every distribution in $\mathcal{B}$. In contrast to expected discounted utility, it is known that in this case the agent's utility will not generally have the recursive property. However, there is a condition on $\mathcal{B}$ called *rectangularity* which is sufficient (and effectively necessary) for the recursivity 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 of distributions $\mathcal{B}$ over S^∞, given an initial value s_0, is then obtained by concatenating the transition probabilities. Therefore, the probabilities associated with a path $(s_1, s_2, s_3, \ldots)$ is given by $\prod_{i=1}^{\infty} p_i$, 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 $\Lambda(s_2)$, etc.

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

$$v(x,s) = \max\left\{u(x,s,y) + \beta(Av)(y,s) : y \in \Phi(x,s)\right\}$$

where $(Av)(y,s) := \min\left\{\int_S v(y,\tilde{s})d\lambda(\tilde{s}) : \lambda \in \Lambda(s)\right\}$ (see Iyengar, 2005). Furthermore, the problem is *well-behaved* in the sense defined at the beginning of this section: the operator $\mathscr{T}$ given by (14), but with the integral replaced by A, has the uniform convergence property and γ is an upper hemi-continuous correspondence 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 7 tells us that this holds if Λ is increasing in the strong set order, and thus we obtain the following proposition. The proof is supplied in the Appendix.

Proposition 11. *Consider a well-behaved optimal control problem where X and S are subsets of $\mathbb{R}$, with X compact and S finite. Suppose that $u(x,s,y)$ is supermodular in* *all three arguments, $\Lambda : S \to \Delta_S$ is increasing in the strong set order, and the graph of $\Phi : X \times S \to X$ is a sublattice. Then the following holds: the value function $v^*(x, s)$ is supermodular; the correspondence $\gamma : X \times S \to \mathbb{R}$, where*

$$\gamma(x, s) := \arg\max \left\{u(x, s, y) + \beta(Av^*)(y, s) : y \in \Phi(x, s)\right\} \tag{17}$$

is sublattice-valued and increasing in the strong set order; and the function $g : X \times S \to \mathbb{R}$, where $g(x, s) := \max \gamma(x, s)$, is an increasing and Borel measurable policy function.

Example H. 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 realised 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

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

where the firm chooses y from the interval $\Phi(x, s) = [\delta x, \delta x + K]$. 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* $\Lambda : S \to \Delta_S$ is increasing with respect to first order stochastic dominance and u is supermodular; the latter property is guaranteed if π is supermodular and c is concave. Proposition 11 goes further by saying that this conclusion remains true if the firm has a maxmin preference, so long as the transition *correspondence* Δ is increasing in the strong set order, with respect to the order induced by first order stochastic dominance.

Appendix

Proof of Proposition 1. We only prove (i); the proof of (ii) analogous. Given that Γ has compact values and ϕ is continuous, function f is well-defined. If Γ is *not* upper supermodular, there exist some x, x' in X and $y \in \Gamma(x)$, $y' \in \Gamma(x')$ such that there are no $z \in \Gamma(x \wedge x')$ and $z' \in \Gamma(x \vee x')$ with $z + z' \geq y + y'$. Define set

$$U := \left\{u \in Y : u \leq v, \text{ for some } v \in \Gamma(x \wedge x') + \Gamma(x \vee x')\right\},$$

which is closed, convex, and downward comprehensive.[30] Given that U is closed and $(y+y') \notin U$, there is some $v^* \notin U$ such that $y + y' \gg v^*$. Moreover, as U is downward comprehensive, we have $\{z \in Y : z \geq v^*\} \cap U = \emptyset$. Let $V := \{z \in Y : z \geq v^*\}$. Since both U and V are non-empty and convex, by the Separating Hyperplane Theorem, there is a non-zero, linear functional ϕ^* such that $\phi^*(v) \geq \phi^*(u)$, for all $v \in V$ and $u \in U$. As $\phi^*(V)$ is bounded below, ϕ^* is also positive.[31] Finally, $y + y' \gg v^*$ implies $\phi^*(y+y') > \phi^*(v^*)$.

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

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

which violates the supermodularity of f. □

Continuation of the proof of Proposition 3. Clearly, Γ_i is well-defined. Take any $p'_i \geq p_i$ and $q' \geq q$. We need to show that, for any $(p_i z_i, z_{-i})$ in $\Gamma_i(p_i, q)$ and $(p'_i z'_i, z'_{-i})$ in $\Gamma_i(p'_i, q')$, there are some $(p'_i y_i, y_{-i}) \in \Gamma_i(p'_i, q)$ and $(p_i y'_i, y'_{-i}) \in \Gamma_i(p_i, q')$ such that

$$(p_i z_i, z_{-i}) + (p'_i z'_i, z'_{-i}) \geq (p'_i y_i, y_{-i}) + (p_i y'_i, y'_{-i}). \tag{A1}$$

By the definition of Γ_i, we have $f(z) \geq q$ and $f(z') \geq q'$. We discuss two cases separately. First, suppose that (i) $z'_i \geq z_i$. Given that

$$p_i z'_i - p_i z_i = p_i(z'_i - z_i) \leq p'_i(z'_i - z_i) = p'_i z'_i - p'_i z_i,$$

we have $p_i z_i + p'_i z'_i \geq p'_i z_i + p_i z'_i$. Choose vectors $y := z$ and $y' := z'$. As $f(y) \geq q$ and $f(y') \geq q'$, by construction, element $(p'_i y_i, y_{-i})$ must belong to $\Gamma_i(p'_i, q)$, while $(p_i y'_i, y'_{-i})$ is in $\Gamma_i(p_i, q')$. Moreover, the two vectors satisfy condition (A1).

Next, we suppose that (ii) $z'_i < z_i$. If (a) $f(z) \geq f(z')$, let $y' := z$ and $y := z'$. By construction, we have $f(y') \geq f(y) \geq q' \geq q$. Hence, element $(p_i y'_i, y'_{-i})$ belongs to $\Gamma_i(p_i, q')$, while $(p'_i y_i, y_{-i})$ is in $\Gamma_i(p'_i, q)$. Moreover, condition (A1) holds trivially. On the

[30] This is the only instance where we use the assumption that Γ is compact-valued. In fact, we only require that Γ satisfies the following property: for any x, x' in X, set $\Gamma(x) + \Gamma(x')$ is closed.

[31] Otherwise, there would be some $w > 0$ such that $\phi^*(w) < 0$ and $(v^* + \lambda w) \in V$, for any $\lambda \geq 0$. However, this would imply that $\phi^*(v^* + \lambda w)$ tends to $-\infty$ as λ tends to ∞.

other hand, if (b) $f(z') > f(z)$, we define $v := z' - (z \wedge z')$. Since $z'_i = (z \wedge z')_i$, it must be that $v_i = 0$. Moreover, the monotonicity of f implies that $f(z \wedge z') \leq f(z) < f(z')$. By continuity of f and the Intermediate Value Theorem, there is some $\lambda \in [0,1]$ such that $f(z \wedge z' + \lambda v) = f(z)$. Since function f is supermodular and i-concave, by Proposition 2 in Quah (2007), it is $\mathcal{C}_i$-supermodular. In particular, this implies that

$$0 \;=\; f(z) - f(z \wedge z' + \lambda v) \;\leq\; f(z \vee z' - \lambda v) - f(z'),$$

and so $f(z \vee z' - \lambda v) \geq f(z') \geq q'$. Let $y := (z \wedge z') + \lambda v$ and $y' := (z \vee z') - \lambda v$, where $(p' y_i, y_{-i}) \in \Gamma_i(p', q)$ and $(p y'_i, y'_{-i}) \in \Gamma_i(p, q')$. Hence, $y + y' = (z \wedge z') + (z \vee z') = z + z'$. Since $y_i = z'_i$ and $y'_i = z_i$, we conclude that condition (A1) holds, which suffices for correspondence Γ_i to be lower supermodular. $\square$

Proof of Proposition 4. Take any x, x' in X, as well as $t \in \Gamma(x)$ and $y' \in \Gamma(x')$. In particular, there are some t', t in T such that $y = f(x,t)$ and $y' = f(x',t')$. Without loss of generality, suppose that $t' \geq t$. We need to show that there is some $z \in \Gamma(x \wedge x')$ and $z' \in \Gamma(x \vee x')$ such that $z + z' \geq y + y'$. Suppose that $f(x',t) \geq f(x',t')$. Then,

$$y + y' \;=\; f(x,t) + f(x',t') \;\leq\; f(x,t) + f(x',t) \;\leq\; f(x \wedge x', t) + f(x \vee x', t),$$

where the last inequality follows from the supermodularity of $f(\cdot, t)$. Since $f(x \wedge x', t) \in \Gamma(x \wedge x')$ and $f(x \vee x', t) \in \Gamma(x \vee x')$, the aforementioned condition is satisfied. An analogous argument using supermodularity of $f(\cdot, t')$ will apply if $f(x,t') \geq f(x,t)$.

Consider the case where $f(x',t) < f(x',t')$ and $f(x,t) > f(x,t')$. We show that

$$f(x,t) + f(x',t') \;\leq\; f(x \wedge x', t) + f(x \vee x', t'), \tag{A2}$$

which establishes the upper supermodularity of Γ since $f(x \wedge x', t) \in \Gamma(x \wedge x')$ and $f(x \vee x', t') \in \Gamma(x \vee x')$. Indeed, since f obeys single crossing differences in (t, x) and $f(x,t) > f(x,t')$, we obtain $f(x \wedge x', t) > f(x \wedge x', t')$. Since $f(x',t) < f(x',t')$, the continuity of f over the convex sublattice X guarantees that there is some $u \in X$ such that $f(u,t) = f(u,t')$, where $u := \lambda x' + (1-\lambda)(x \wedge x')$, for $\lambda \in (0,1)$. Notice that

$$\begin{aligned} f(x',t') - f(x' \wedge x', t) \;&=\; \big[f(x',t') - f(u,t')\big] + \big[f(u,t) - f(x \wedge x', t)\big] \\ &\leq\; \big[f(x \vee x', t') - f(v,t')\big] + \big[f(v,t) - f(x,t)\big] \;\leq\; f(x \vee x', t') - f(x,t), \end{aligned}$$

where $v := \lambda(x \vee x') + (1-\lambda)x$. The first inequality follows from supermodularity of f over X. Moreover, by single crossing differences, if $f(u,t') = f(u,t)$ then $f(v,t') \geq f(v,t)$, which implies the second inequality. Therefore, condition (A2) holds.[32] $\square$

Given that the proof of Proposition 5 requires Proposition 6, we find it convenient to prove the former first.

Proof of Proposition 6. Suppose $v(x) - v(x \wedge x') \geq (>)\ 0$, for some x, x' in X_Z. There is some t, t' in T such that $v(x) = f(x,t)$ and $v(x') = f(x',t')$. Since Z is a sublattice of $X \times T$, we have $(t \wedge t') \in Z(x \wedge x')$, and thus $v(x \wedge x') \geq f(x \wedge x', t \wedge t')$. Therefore,

$$f(x,t) - f(x \wedge x', t \wedge t') \;\geq\; v(x) - v(x \wedge x') \;\geq (>)\ 0.$$

Given that f is quasisupermodular, this inequality implies that

$$f(x \vee x', t \vee t') - f(x', t') \;\geq (>)\ 0.$$

To complete the proof, recall that $v(x \vee x') \geq f(x \vee x', t \vee t')$ and $v(x') = f(x', t')$. This allows us to conclude that $v(x \vee x') - v(x') \geq (>)\ 0$. $\square$

Proof of Proposition 5. We define the function $f_{-m} : X \times \left(\times_{j=1}^{m-1} T_j\right) \to \mathbb{R}$ by

$$f_{-m}(x, t_{-m}) = \max\left\{f(x, t_{-m}, t_m) : t_m \in T_m\right\}.$$

This function is well-defined since, in particular, T_m is compact and f is continuous in t_m. We claim that, like the function f itself, f_{-m} has properties (i) – (iv) as well. That f_{-m} is continuous in (x, t_{-m}) follows from the joint continuity of f in (x,t). It is also supermodular in x since f is supermodular in x and has single crossing differences in (t_m, x) (by Proposition 4). Furthermore it is quasisupermodular in t_{-m} because f is quasisupermodular in t (by Proposition 6). It remains for us to show that f_{-m} has property (iv), i.e., f_{-m} has single crossing differences in (t_{-m}, x). This claim can be established by a proof similar to the proof of Proposition 6. Indeed, suppose $t'_{-m} \geq t_{-m}$ and for some $x \in X$, we have $f_{-m}(x, t'_{-m}) - f_{-m}(x, t_{-m}) \geq (>)\ 0$. We want to show

[32] Notice that the proof only requires that $f(x,t)$ obey *weak single crossing differences in* (t,x), i.e., if $t' \geq t$ and $f(x,t') \geq f(x,t)$ then $f(x',t') \geq f(x',t)$, for all $x' \geq x$. It does not require that $f(x,t') > f(x,t)$ implies $f(x',t') > f(x',t)$, for all $x' \geq x$.

that $f_{-m}(x', t'_{-m}) - f_{-m}(x', t_{-m}) \geq (>) \; 0$, if $x' \geq x$. There is $t'_m \in T_m$ that satisfies $f_{-m}(x, t'_{-m}) = f(x, t'_{-m}, t'_m)$ and $t_m \in T_m$ that satisfies $f_{-m}(x', t_{-m}) = f(x, t_{-m}, t_m)$. By the definition of f_{-m}, we have $f_{-m}(x, t_{-m}) \geq f(x, t_{-m}, t_m \wedge t'_m)$ and therefore

$$f(x, t'_{-m}, t'_m) - f(x, t'_{-m}, t_m \wedge t'_m) \;\geq\; f_{-m}(x, t'_{-m}) - f_{-m}(x, t_{-m}) \;\geq (>) \; 0.$$

Since f is quasisupermodular in t, this inequality implies that

$$f(x, t'_{-m}, t_m \vee t'_m) - f(x, t_{-m}, t_m) \;\geq (>) \; 0.$$

Since $(t'_{-m}, t_m \vee t'_m) \geq (t_{-m}, t_m)$ and f has single crossing differences in (t, x), we obtain

$$f(x', t'_{-m}, t_m \vee t'_m) - f(x', t_{-m}, t_m) \;\geq (>) \; 0.$$

It follows that $f_{-m}(x', t'_{-m}) - f_{-m}(x', t_{-m}) \geq (>) \; 0$. Since this argument can be repeated, the function $f_{-\{m-1,m\}} : X \times \left(\times_{j=1}^{m-2} T_j\right) \to \mathbb{R}$ given by

$$f_{-\{m-1,m\}}(x, t_{-\{m-1,m\}}) = \max\left\{f_{-m}(x, t_{-\{m-1,m\}}, t_{m-1}) : t_{m-1} \in T_{m-1}\right\}$$

is well-defined and has properties (i) – (iv). Repeating this argument across all the dimensions of T, we conclude that $v(x) = f_{-\{1,2,\cdots,m\}}(x)$ is a supermodular function. □

Continuation of the proof of Proposition 7. To show that Γ is upper submodular, take any $x' \geq x$ in X, $t' \geq t$ in T, and let $y \in \Gamma(x, t')$ and $y' \in \Gamma(x', t)$. By definition of Γ, there is some $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$ such that $y_i = \delta_i(x)\lambda'(s_i)$ and $y'_i = \delta_i(x')\lambda(s_i)$, for all $i = 1, \ldots, \ell$. Since the correspondence Λ is increasing in the strong set order, we obtain $(\lambda \wedge \lambda') \in \Lambda(t)$ and $(\lambda \vee \lambda') \in \Lambda(t')$, where $(\lambda \wedge \lambda')(s_i) = \max\left\{\lambda(s_i), \lambda'(s_i)\right\}$ and $(\lambda \vee \lambda')(s_i) = \min\left\{\lambda(s_i), \lambda'(s_i)\right\}$, for all $i = 1, \ldots, n$. By the supermodularity of f, $\delta_i(x') \geq \delta_i(x)$, and so

$$\delta_i(x')\big[(\lambda \vee \lambda')(s_i) - \lambda(s_i)\big] \;\leq\; \delta_i(x)\big[\lambda'(s_i) - (\lambda \wedge \lambda')(s_i)\big]. \tag{A3}$$

Define the vectors z, z' in $\mathbb{R}^n$ by $z_i := \delta_i(x)(\lambda \wedge \lambda')(s)$, and $z'_i(s) := \delta_i(x')(\lambda \vee \lambda')(s)$, for $i = 1, \ldots, \ell$. Clearly, $z \in \Gamma_n(x, t)$ and $z' \in \Gamma_n(x', t')$, while (A3) implies $z + z' \leq y + y'$. Thus, Γ is upper submodular.

To prove that Γ is lower submodular, take any $x' \geq x$ in X, $t' \geq t$ in T, and let $z \in \Gamma(x, t)$ and $z' \in \Gamma(x', t')$. By definition, there are distributions $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$

such that $z_i = \delta_i(x)\lambda(s_i)$ and $z'_i = \delta_i(x')\lambda'(s_i)$, for $i = 1, \ldots, \ell$. Since Λ is monotone, $(\lambda \wedge \lambda') \in \Lambda(t)$ and $(\lambda \vee \lambda') \in \Lambda(t')$. Since $\delta_i(x') \geq \delta_i(x)$ we obtain

$$\delta_i(x')\big[(\lambda \wedge \lambda')(s_i) - \lambda'(s_i)\big] \;\geq\; \delta_i(x)\big[\lambda(s_i) - (\lambda \vee \lambda')(s_i)\big], \tag{A4}$$

for all i. Define the vectors y and y' in $\mathbb{R}^n$ by $y_i := \delta_i(x)(\lambda \vee \lambda')(s_i)$ and $y'_i := \delta_i(x')(\lambda \wedge \lambda')(s_i)$ for $i = 1, \ldots, \ell$. Since $y \in \Gamma_n(x, t')$ and $y' \in \Gamma_n(x', t)$, we obtain $y + y' \geq z + z'$ from (A4). We conclude that Γ is lower submodular.

Proof in the case where S is an interval (as in Remark). Suppose $S = [a, b]$; for each N, we choose $\{s_i^N\}_{i=1}^N$ such that $a = s_0^N < s_1^N < \ldots < s_{N-1}^N < s_N^N = b$, with the mesh approaching 0 as $N \to \infty$. Since at each (x, t), $f(x, \cdot)$ is Riemann-Stieltjes integrable with respect to $\lambda \in \Lambda(t)$, we obtain

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

for all $\lambda \in \Lambda(t)$. This guarantees that $\lim_{N\to\infty} v_N(x, t) = v(x, t)$ for all (x, t), where

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

We know, from the case where S is finite, that $v_N : X \times T \to \mathbb{R}$ is a supermodular function. Since supermodularity is preserved by pointwise convergence, v is also supermodular. An analogous argument guarantees that w is a supermodular function. □

Proof of Proposition 8. First, denote $S = \{s_i\}_{i=1}^{\ell+1}$ such that $s_1 < s_2 \ldots < s_{\ell+1}$. From the proof of Proposition 7, we know that, for any $x \in X$ and $\lambda \in \Delta_S$,

$$\int_S f(x, s)d\lambda(s) = f(x, s_{\ell+1}) - \sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i),$$

where $\delta_i(x) := \big[f(x, s_{i+1}) - f(x, s_i)\big]$, for $i = 1, \ldots, \ell$. Therefore,

$$v(x, t) = f(x, s_{\ell+1}) - \max\left\{\sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i) - c(\lambda, t) : \lambda \in \Lambda(t)\right\}, \tag{A5}$$

which is supermodular whenever $\bar{v}(x, t) := \max\big\{\sum_{i=1}^{\ell} \delta_i(x)\lambda(s_i) - c(\lambda, t) : \lambda \in \Lambda(t)\big\}$ is submodular. By the Main Theorem (∗), this holds if the correspondence $\bar{\Gamma} : X \times T \to \mathbb{R}^\ell$,

$$\bar{\Gamma}(x, t) := \Big\{(a, b) \in \mathbb{R}^\ell \times \mathbb{R} : a_i = \delta_i(x)\lambda(s_i), \text{ for } i = 1, \ldots, \ell, \\ \text{and } b = -c(\lambda, t), \text{ for some } \lambda \in \triangle_S\Big\},$$

is lower submodular. Take any $x' \geq x$ in X and $t' \geq t$ in T, as well as $(a', b') \in \bar{\Gamma}(x', t')$ and $(a, b) \in \bar{\Gamma}(x, t)$. Then there is λ and λ' in Δ_S such that $a'_i = \delta_i(x')\lambda(s_i)$, and $a_i = \delta_i(x)\lambda(s_i)$, for $i = 1, 2, \ldots, \ell$, with $b' = -c(\lambda', t')$ and $b = -c(\lambda, t)$. It follows from the proof of Proposition 7 (specifically from the second part) that

$$d + d' \geq a + a'$$

where $d_i = \delta_i(x')(\lambda \wedge \lambda')(s_i)$ and $d'_i = \delta_i(x)(\lambda \vee \lambda')(s_i)$, for $i = 1, 2, \ldots, \ell$. Note that $\big(d', -c(\lambda \vee \lambda', t')\big) \in \bar{\Gamma}(x, t')$ and $\big(d, -c(\lambda \wedge \lambda', t)\big) \in \bar{\Gamma}(x', t)$. We obtain

$$\big(d', -c(\lambda \vee \lambda', t')\big) + \big(d, -c(\lambda \wedge \lambda', t)\big) \;\geq \big(a, -c(\lambda, t)\big) + \big(a', -c(\lambda', t')\big)$$

and thus the lower submodularity of $\bar{\Gamma}$, provided c is submodular. □

Proof of Proposition 9. To show that $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular in (λ, t) it suffices to show that it is submodular in λ and has decreasing differences in (λ, t).

We first show that $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular in λ. We denote by $\pi_i^\wedge$ and $\pi_i^\vee$ the probability of state i for the distributions $(\lambda \wedge \lambda')$ and $(\lambda \vee \lambda')$ respectively. It suffices to show that, for all $i = 1, 2, \ldots, \ell + 1$,

$$\pi_i \ln \pi_i + \pi'_i \ln \pi'_i - (\pi_i + \pi'_i) \ln \pi^*_i(t) \;\geq\; \pi_i^\wedge \ln \pi_i^\wedge + \pi_i^\vee \ln \pi_i^\vee - (\pi_i^\wedge + \pi_i^\vee) \ln \pi^*_i(t)$$

Clearly this claim is true for $i = 1$. Assuming that it is true for some $i = k$, we now show that it holds for $i = k + 1$. With no loss of generality, suppose $(\lambda \wedge \lambda')(s_k) = \lambda(s_k)$ and $(\lambda \vee \lambda')(s_k) = \lambda'(s_k)$. We need to consider two cases. In the first case,

$$\pi'_{k+1} + \lambda'(s_k) \;\leq\; \pi_{k+1} + \lambda(s_k)$$

so that $(\lambda \wedge \lambda')(s_{k+1}) = \lambda(s_{k+1})$ and $(\lambda \vee \lambda')(s_{k+1}) = \lambda'(s_{k+1})$. Then $\pi^\wedge_{k+1} = \pi_{k+1}$ and $\pi^\vee_{k+1} = \pi'_{k+1}$ and it is obvious that

$$\pi_{k+1} \ln \pi_{k+1} + \pi'_{k+1} \ln \pi'_{k+1} - \big(\pi_{k+1} + \pi'_{k+1}\big) \ln \pi^*_{k+1}(t) \;=\; \pi^\wedge_{k+1} \ln \pi^\wedge_{k+1} + \pi^\vee_{k+1} \ln \pi^\vee_{k+1} - \big(\pi^\wedge_{k+1} + \pi^\vee_{k+1}\big) \ln \pi^*_{k+1}(t)$$

In the second case,

$$\pi'_{k+1} + \lambda'(s_k) \;>\; \pi_{k+1} + \lambda(s_k),$$

so that $(\lambda \wedge \lambda')(s_{k+1}) = \lambda'(s_{k+1})$ and $(\lambda \vee \lambda')(s_{k+1}) = \lambda(s_{k+1})$. Let $d := \lambda(s_k) - \lambda'(s_k)$ and note that $0 \leq d < \pi'_{k+1} - \pi_{k+1}$. Since $\pi^{\wedge}_{k+1} = \pi'_{k+1} - d$ and $\pi^{\vee}_{k+1} = \pi_{k+1} + d$,

$$\pi^{\wedge}_{k+1} \ln \pi^{\wedge}_{k+1} + \pi^{\vee}_{k+1} \ln \pi^{\vee}_{k+1} - \big(\pi^{\wedge}_{k+1} + \pi^{\vee}_{k+1}\big) \ln \pi^{*}_{k+1}(t) \;=\; (\pi'_{k+1} - d) \ln(\pi'_{k+1} - d) + (\pi_{k+1} + d) \ln(\pi_{k+1} + d) - \big(\pi_{k+1} + \pi'_{k+1}\big) \ln \pi^{*}_{k+1}(t)$$

This is less than $\pi_{k+1} \ln \pi_{k+1} + \pi'_{k+1} \ln \pi'_{k+1} - \big(\pi_{k+1} + \pi'_{k+1}\big) \ln \pi^{*}_{k+1}(t)$ because the map from z to $z \ln z$ is convex.

To show that $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ has decreasing differences in (λ, t), we first note that

$$R\big(\lambda' \| \lambda^*(\cdot, t)\big) - R\big(\lambda \| \lambda^*(\cdot, t)\big) = \sum_i \pi'_i \ln \pi'_i - \sum_i \pi_i \ln \pi_i + \sum_i (\pi_i - \pi'_i) \ln \pi^*_i(t)$$

Therefore, for $t' \geq t$,

$$\Big[R\big(\lambda' \| \lambda^*(\cdot, t')\big) - R\big(\lambda \| \lambda^*(\cdot, t')\big)\Big] - \Big[R\big(\lambda' \| \lambda^*(\cdot, t)\big) - R\big(\lambda \| \lambda^*(\cdot, t)\big)\Big] \;=\; \sum_i \pi_i \big(\ln \pi^*_i(t') - \ln \pi^*_i(t)\big) - \sum_i \pi'_i \big(\ln \pi^*_i(t') - \ln \pi^*_i(t)\big).$$

This expression is non-positive since $\ln \pi^*_i(t') - \ln \pi^*_i(t)$ is an increasing function of i (because $\lambda^*(t)$ is increasing with t with respect to the monotone likelihood ratio order) and λ' first order stochastically dominates λ. $\square$

Proof of Proposition 10. Before we proceed with our argument, we need to introduce one auxiliary result. Our proof uses the following extension of Proposition 5.[33]

Lemma .1. *Let X be a convex sublattice of $\mathbb{R}^n$, S a convex subset of $\mathbb{R}^k$, and $T = \times_{i=1}^m T_i$, where T_i are compact subsets of $\mathbb{R}$. Suppose that the function $f : X \times S \times T \to \mathbb{R}$ has the following properties:* (i) *it is jointly continuous in (x, s, t);* (ii) *supermodular in x;* (iii) *has increasing differences in (x, s);* (iv) *quasisupermodular in t; and* (iv) *obeys single crossing differences in $(t, (x, s))$. Then, function $v(x, s) := \max\big\{f(x, s, t) : t \in T\big\}$ is jointly continuous in (x, s), supermodular in x and has increasing differences in (x, s).*

Proof. The joint continuity of v in (x, s) follows immediately from the joint continuity of u in (x, s, t). Its supermodularity in x follows from Proposition 5. It remains to show

[33] This extension is loosely analogous to the extension of the Topkis Value Theorem developed and used in Hopenhayn and Prescott (1992).

that v has increasing differences in (x,s). Suppose $x' \geq x$ and $s' \geq s$. We claim that

$$v(x', s') - v(x, s') \geq v(x', s) - v(x, s). \tag{A6}$$

To do this we first define the function $h : [0,1] \times [0,1] \times T \to \mathbb{R}$ where

$$h\big(\beta_1, \beta_2, t\big) \;:=\; f\big((1-\beta_1)x + \beta_1 x', (1-\beta_2)x + \beta_2 x', t\big).$$

This function is well-defined since X and S are both convex sets. Furthermore, h has the following properties: it is continuous in (β_1, β_2, t) — because f is continuous; it is supermodular in (β_1, β_2) — because f has increasing differences in (x,s); it is quasisupermodular in t — because f is quasisupermodular in t; and it obeys single crossing differences in $\big(t, (\beta_1, \beta_2)\big)$ — because f obeys single crossing differences in $\big(t, (x,s)\big)$. Applying Proposition 5 to h, we conclude that $H(\beta_1, \beta_2) = \max_{t \in T} h(\beta_1, \beta_2, t)$ is a supermodular function. In particular, we obtain $H(1,1) - H(0,1) \geq H(1,0) - H(0,0)$, which is equivalent to (A6). This completes the proof of the lemma. □

To prove Proposition 10, we first claim that if $w(x,s)$ is decreasing in x, supermodular in x, and has increasing differences in (x,s), then $\mathscr{T}w$ will have the same properties. The last two properties follow immediately from the Lemma so long as we can show that $\alpha(x,s,y) := u(x,s,y) + \beta \int_S w(y,\tilde{s}) d\lambda(\tilde{s}, s)$ is supermodular in x, quasisupermodular in y, has increasing differences in (x,s) and has single crossing differences in $\big(y,(x,s)\big)$. Since u is supermodular in x, so is α. Furthermore, α is supermodular in y since both u and w are supermodular in y. It remains for us to show that α obeys single crossing differences in $\big(y,(x,s)\big)$. Suppose $\alpha(x,s,y') - \alpha(x,s,y) \geq (>)\, 0$ for some $y' \geq y$. Since $w(y,s)$ is decreasing in y, for all $s \in S$, this implies that $u(x,s,y') \geq (>) u(x,s,y)$. It follows from the conditional increasing differences property on u that, for any $(x', s') > (x,s)$,

$$u(x', s', y') - u(x', s', y) \;\geq\; u(x,s,y') - u(x,s,y) \;\geq (>)\; 0.$$

Furthermore, $\int_S \big[w(y', \tilde{s}) - w(y, \tilde{s})\big]\, d\lambda(\tilde{s}, s') \geq \int_S \big[w(y', \tilde{s}) - w(y, \tilde{s})\big]\, d\lambda(\tilde{s}, s)$ — because $w(y,s)$ has increasing differences in (y,s) and $\lambda(\cdot, s)$ is increasing with respect to first order stochastic dominance. We conclude that

$$\alpha(x', s', y') - \alpha(x', s', y) \;\geq\; \alpha(x,s,y') - \alpha(x,s,y).$$

Therefore, α obeys single crossing differences in $\big(y,(x,s)\big)$. We conclude that $(\mathscr{T}w)$ is supermodular in x and has increasing differences in (x,s). Furthermore, it is decreasing in x since u is decreasing in x. So we have shown that $\big[w(y',\tilde{s}) - w(y,\tilde{s})\big]$ inherits the properties of w. This implies that $v^*(x,s)$, which is the limit of $\mathscr{T}^n w$ as $n \to \infty$ is also decreasing in x, supermodular in x, and has increasing differences in (x,s).

By setting $w = v^*$, we know that the objective function in (15) is supermodular in y and has single crossing differences in $\big(y,(x,s)\big)$. The Milgrom-Shannon MCS Theorem guarantees that $\gamma(x,s)$ is a sublattice and it increases with (x,s) in the strong set order. Since $\gamma(x,s)$ is also non-empty and compact (because the problem is assumed to be well-behaved), it follows that $g(x,s)$ exists and is increasing in (x,s). By the argument in Hopenhayn and Prescott, 1992) it is also Borel measurable. □

Proof of Proposition 11. Let $w : X \times S \to \mathbb{R}$ be any continuous function. Since the problem is well-behaved we know that the function $(\mathscr{T}w)$, given by

$$\big(\mathscr{T}w\big)(x,s) = \max\big\{u(x,s,y) + \beta(Aw)(y,s) : y \in \Phi(x,s)\big\},$$

is another continuous function on $X \times S$ and $\mathscr{T}^n w$ converges uniformly to v^* as $n \to \infty$. By Proposition 7, Aw is a supermodular function, and thus $u(x,s,y) + \beta(Aw)(y,s)$ is a supermodular function of (x,s,y). By the Topkis Value Theorem, function $(\mathscr{T}w)$ is supermodular. Taking limits as $n \to \infty$, we conclude that v^* is a supermodular function and hence the objective in (17) is supermodular in (x,s,y). Furthermore, since the graph of Φ is a sublattice, $\Phi(x,s)$ is sublattice of X and it increases with (x,s) in the strong set order. Applying a stronger version of the Topkis MCS Theorem that allows for increasing constraint sets (see Topkis, 1978), we conclude that $\gamma(x,s)$ is a sublattice and it increases with (x,s) in the strong set order. Since the problem is well-behaved, $\gamma(x,s)$ is non-empty and compact. Therefore, $g(x,s)$ exists and is increasing, as well as Borel measurable, which follows from Hopenhayn and Prescott (1992). □

References

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

Amir, R. (1996): "Cournot oligopoly and the theory of supermodular games," *Games and Economic Behavior*, 15(2), 132–148.

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.

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.

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

Milgrom, P. (2004): *Putting Auction Theory to Work*. Cambridge University Press.

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.