ECONOMICS

Working Paper Series

No. 01-2021

Comparative statics with linear objectives: normal demand, monotone marginal costs, and ranking multi-prior beliefs

Pawel Dziewulski
University of Sussex
p.k.dziewulski@sussex.ac.uk

John K.-H. Quah
John Hopkins University
National University of Singapore
john.quah@jhu.edu

Abstract: We formulate a set order on constraint sets $C \subset R^\ell$ which guarantee that argmin $\{\phi(x) : x \in C\}$ increases in the product order as C increases in the set order, for all linear functions $\phi: R^\ell \to R$. Using this result, we characterize the utility/production functions that lead to normal demand; we also show that this very same class of production functions have marginal costs that increase with factor prices. In the context of decision-making under uncertainty, our new set order leads to natural generalizations of first order stochastic dominance in multi-prior models.

JEL classification: C61, D21, D24

Key words: parallelogram property, increasing differences, ambiguity, first order stochastic dominance, normal demand, marginal costs

Comparative statics with linear objectives: normal demand, monotone marginal costs, and ranking multi-prior beliefs*

Paweł Dziewulski† John K.-H. Quah‡

February, 2021

Abstract

We formulate a set order on constraint sets $C \subset \mathbb{R}^\ell$ which guarantee that $\operatorname{argmin}\{\phi(x) : x \in C\}$ increases in the product order as C increases in the set order, for all linear functions $\phi : \mathbb{R}^\ell \to \mathbb{R}$. Using this result, we characterize the utility/production functions that lead to normal demand; we also show that this very same class of production functions have marginal costs that increase with factor prices. In the context of decision-making under uncertainty, our new set order leads to natural generalizations of first order stochastic dominance in multi-prior models.

Keywords: parallelogram property, increasing differences, ambiguity, first order stochastic dominance, normal demand, marginal costs

JEL Classification: C61, D21, D24

* We would like to thank Eddie Dekel, Ludvig Sinander, Tomasz Strzalecki, Bruno Strulovici, and Takashi Ui for helpful comments. We are also grateful to the participants of the seminars at Northwestern University, Georgetown University, Université Libre de Bruxelles, University of Maryland, HKU-CUHK-HKUST (online) and audiences attending 2019 CETC in Montreal, 2019 RUD in Paris, 2019 EWET in Berlin, 2019 NSF/NBER/CEME Mathematical Economics Conference in Berkeley, 2019 Econometric Society Winter Meeting in Rotterdam, and 2020 SET organised by University of Warwick (online) for helpful comments.

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

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

1 Introduction

Suppose a firm produces a single product using ℓ inputs and has the production function $f : \mathbb{R}^\ell_+ \to \mathbb{R}_+$. We denote by C the set of input/factor combinations that could produce output q, i.e., $C = \{x \in \mathbb{R}^\ell_+ : f(x) \geq q\}$, and by C' the corresponding set of input combinations that produce output $q' > q$. Let $p = (p_1, p_2, \ldots, p_\ell)$ be the price of the ℓ factors and suppose x is the conditional factor demand at q, i.e. $x \in \operatorname{argmin}\{p \cdot \tilde{x} : \tilde{x} \in C\}$. Question: when can we guarantee that the firm's conditional factor demand is *normal* in the sense that there is $x' \in \operatorname{argmin}\{p \cdot \tilde{x} : \tilde{x} \in C'\}$ such that $x' \geq x$ (in other words, the demand for every factor has increased)?

While this seems like a very basic question, it is not one for which the existing tools of monotone comparative statics gives a wholly satisfactory answer. A canonical result in monotone comparative statics says the following: suppose that $\phi : \mathbb{R}^\ell_+ \to \mathbb{R}$ is a submodular function (or its ordinal analog, a quasisubmodular function) and C' dominates C in the sense of the strong set order; then for any $x \in \operatorname{argmin}\{\phi(\tilde{x}) : \tilde{x} \in C\}$ there is $x' \in \operatorname{argmin}\{\phi(\tilde{x}) : \tilde{x} \in C'\}$ such that $x \leq x'$ (see Topkis, 1978 and Milgrom and Shannon, 1994).[1] If we apply this result to our problem, we see that the map from x to $p \cdot x$ is indeed submodular, but C' and C are typically not ordered by the strong set order. The strong set order requires that for any $x \in C$ and $x' \in C'$, we have $x \vee x' \in C'$ and $x \wedge x' \in C$.[2] Provided the production function f is increasing, $x \vee x'$ is indeed in C', but one simply *cannot* guarantee that $x \wedge x'$ is in C so long as f is strictly increasing.

Quah (2007) goes some way towards addressing the question we posed by considering a class of objective functions with a property stronger than submodularity (but weaker than linearity), which then allows for the recovery of some plausible conditions on the production function f that guarantee normality. However, the conditions recovered by that approach are not necessary for normality.

In this paper, we directly answer the question we pose by identifying the relationship between C and C' that characterizes normality, for all *linear* objective functions ϕ. We call this condition the *parallelogram property*; when C and C' are convex, the property says the following: for any $x \in C$ and $x' \in C'$, there is $y \in C$ and $y' \in C'$ such that $x \leq y'$,

[1] In fact, the conclusion is stronger, namely, that $\operatorname{argmin}\{\phi(\tilde{x}) : \tilde{x} \in C'\}$ also dominates $\operatorname{argmin}\{\phi(\tilde{x}) : \tilde{x} \in C\}$ in the strong set order.

[2] For any $x, x' \in \mathbb{R}^\ell$, we denote $(x \wedge x')_i = \min\{x_i, x'_i\}$ and $(x \vee x')_i = \max\{x_i, x'_i\}$, for all $i = 1, \ldots, \ell$.

$y \leq x'$, and the four points form a parallelogram, i.e., $x + x' = y + y'$.[3] (Notice that the strong set order is the special case of this property where $y = x \wedge x'$ and $y' = x \vee x'$.) We also show that there is a more nuanced version of the parallelogram property that is necessary and sufficient to guarantee that the demand for factor i (but not necessarily the other factors) is normal, in the sense of being increasing with output.

Obviously, the parallelogram property encompasses all previously known sufficient conditions for normality. It is also more general than the characterizations of normality obtained in Alarie et al. (1990) and Bilancini and Boncinelli (2010), which require the utility/production function to be strictly quasiconcave and twice-differentiable, with demand being unique, strictly positive for all commodities, and a smooth function of prices. Our formulation requires none of these background assumptions.

As an important application of our new characterization, we obtain, in a fully general setting, the following equivalence: the conditional factor demand for a factor i is normal if and only if an increase in the price of i leads to higher marginal cost. This equivalence was first observed by Fisher (1990), but Fisher's proof required that factor demand be unique and differentiable; we establish this equivalence allowing for production functions that are neither smooth nor quasiconcave. This result in turn allows us to conclude that a necessary and sufficient condition for a lower factor price to lead to higher output is for that factor to be normal. This characterization may be surprising at first blush, since it appears intuitive that lowering the price of a factor would stimulate output, but that is not generally true. What *is* generally true is that lowering the price of a factor raises the demand for that factor, but further restrictions on the production function (specifically the parallelogram property) is needed to guarantee that the firm chooses to produce more.

The applicability of our basic theorem goes beyond comparative statics in producer and consumer theory. It can also be used to formulate an appropriate notion of first order stochastic dominance in multi-prior models; we do this for the maxmin model of Gilboa and Schmeidler (1989) as well as for the variational preference model of Maccheroni et al. (2006), a special case of which is the multiplier preference model (see Sargent and Hansen, 2001 and Strzalecki, 2011a). While there is considerable work on comparative ambiguity and its impact on choice (see, e.g, Gollier, 2011), we find very little discussion of first

[3] This requirement only has a bite when $x \not\leq x'$; otherwise, we can simply let $y = x$ and $y' = x'$.

order stochastic dominance in multi-prior models; one exception that we are aware of is Ui (2015), but that analysis pertains to the specific choice environment of global games.

To understand our contribution, consider an agent who has to take an action under uncertainty. Suppose that the agent's payoff is $g(x, s)$ if $x \in X \subseteq \mathbb{R}$ is the chosen action and $s \in S \subseteq \mathbb{R}$ is the realized state. If the agent maximizes expected utility, then the utility of action x is $f(x, t) := \int g(x, s) d\lambda(s, t)$, where $t \in T \subseteq \mathbb{R}$ parameterizes the distribution function $\lambda(\cdot, t)$ over S. Now suppose that g is such that the marginal payoff of a higher action increases with s, i.e., the function g is supermodular; then perhaps the *expected* marginal payoff of a higher action should be greater when higher states are more likely. This intuition is correct: f is supermodular (in (x, t)) for any supermodular function g if and only if $\lambda(\cdot, t)$ increases with t with respect to first order stochastic dominance, i.e., if $t' > t$ then $\lambda(\cdot, t')$ first order stochastically dominates $\lambda(\cdot, t)$, which we denote by $\lambda(\cdot, t') \succeq \lambda(\cdot, t)$.[4] The supermodularity of f in turn implies that the agent's optimal action, i.e., $\operatorname{argmax}\{f(x, t) : x \in X\}$, increases with t.

Suppose that instead of maximizing expected utility, the agent is ambiguity averse and has maxmin preferences, so that the ex-ante utility of action x is

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

where $\Lambda(t)$ denotes a convex set of probability distributions over S parameterized by t. Assuming that g is supermodular, what conditions on the correspondence Λ are necessary and sufficient for the supermodulatiy of f? In other words, how do we compare *sets of distributions* in a way that generalizes first order stochastic dominance?

An application of our fundamental theorem to this question yields the parallelogram property as the appropriate set-generalization of first order stochastic dominance. In this context, the parallelogram property says the following: if $t' > t$, then for any $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$, there is some $\mu \in \Lambda(t)$ and $\mu' \in \Lambda(t')$ such that $\lambda' \succeq \mu$, $\mu' \succeq \lambda$ and $\frac{1}{2}\lambda + \frac{1}{2}\lambda' = \frac{1}{2}\mu + \frac{1}{2}\mu'$. A simple example where this property holds is when $\Lambda(t) = \{\lambda \in \triangle_S : \underline{\nu}(\cdot, t) \succeq \lambda \succeq \bar{\nu}(\cdot, t)\}$, where $\bar{\nu}(\cdot, t) \succeq \underline{\nu}(\cdot, t)$ and both $\bar{\nu}(\cdot, t)$ and $\underline{\nu}(\cdot, t)$ are increasing in t.[5] More examples are provided in the text.

[4] For a proof of this claim, as well as applications where g is supermodular, see Section 4.1.

[5] In this case, for any $\lambda \in \Lambda(t)$ and $\lambda' \in \Lambda(t')$, the parallelogram property holds because we can choose $\mu = \lambda \wedge \lambda'$ and $\mu' = \lambda \vee \lambda'$.

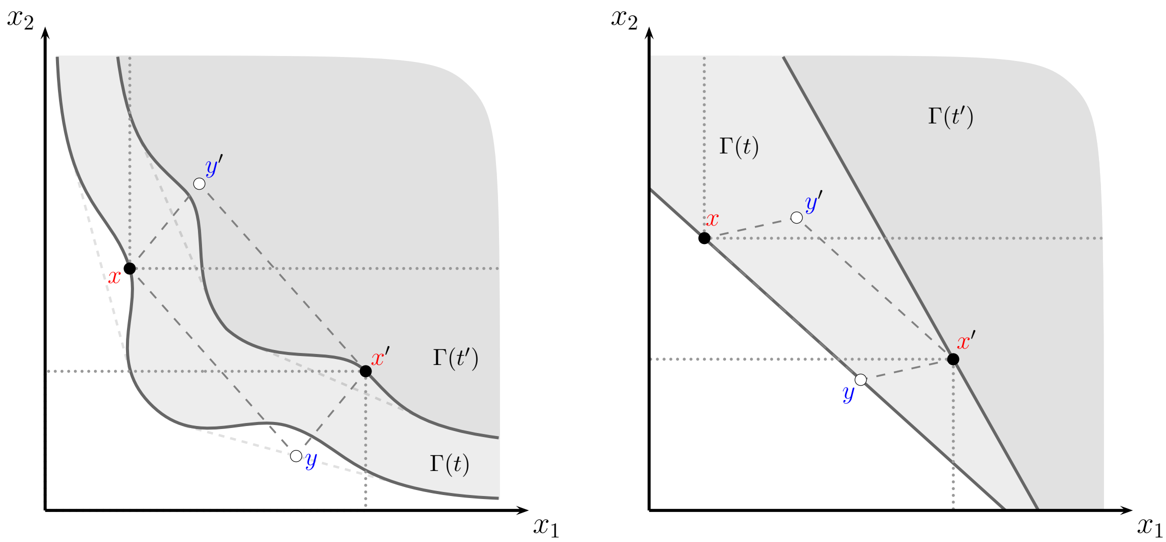
It is noteworthy that the parallelogram property is just one of at least two reasonable set-generalizations of first order stochastic dominance. A related and perhaps more basic problem is the following: what property on the correspondence Λ guarantees that $\min\left\{\int_S u(s)d\lambda : \lambda \in \Lambda(t)\right\}$ is increasing in t, for any increasing function $u : S \to \mathbb{R}$? We show that this leads to a property on Λ that is weaker than the parallelogram property; thus, in multi-prior models, there is an interesting distinction between stochastic dominance that guarantees monotone utility and stochastic dominance that guarantees monotone comparative statics.

Organization of the paper. The basic comparative statics result is set out in Section 2. The rest of the paper is devoted to the two applications. Section 3 applies the basic theorem to guarantee normality in conditional factor demand and to guarantee that marginal cost increases with factor prices; the closely related issue of normal Marshallian demand is covered in Section 3.4. Section 4 applies our result to comparative statics problems with ambiguity averse agents and formulates first order stochastic dominance in multi-prior models; Section 4.1 covers the maxmin model and Section 4.2 covers the variational and multiplier preference models. Readers who are only interested in the applications to ambiguity averse decision-making can skip Section 3 (on normal demand). There is an Appendix containing the more elaborate proofs and also an Online supplement with supplementary material. In particular, the Online supplement contains results on monotone decision rules for ambiguity averse agents, generalizing known results (see Hopenhayn and Prescott (1992)) for agents maximizing discounted expected utility.

2 Parallelogram property

A *partial order* $\geq_X$ over a set X is a reflexive, transitive, and antisymmetric binary relation. A partially ordered set, or a *poset*, is a pair $(X, \geq_X)$ consisting of a set X and a partial order $\geq_X$. Whenever it causes no confusion, we denote $(X, \geq_X)$ with X.

Most of our analysis is carried out in the ℓ-dimensional Euclidean space $\mathbb{R}^\ell$. For any vector $x \in \mathbb{R}^\ell$, its i'th entry is denoted by x_i; for any set $K \subseteq \{1, 2, \ldots, \ell\}$, let $x_K := (x_i)_{i \in K}$ be the sub-vector of entries in x that belong to K. Thus, we can write x as (x_K, x_{-K}), where $x_{-K} := (x_i)_{i \notin K}$. The product partial order $\geq$ on $\mathbb{R}^\ell$ is defined as



(a) Correspondence Γ satisfies parallelogram property for $K = \{1, 2\}$.

(b) Parallelogram property is violated for $K = \{2\}$ and $K = \{1, 2\}$.

Figure 1: Geometric interpretation of parallelogram property.

follows: for any x, $x' \in \mathbb{R}^\ell$, we have $x' \geq x$ if $x'_i \geq x_i$, for all $i = 1, 2, \ldots, \ell$. The relation is said to be *strict*, and denoted by $x' > x$, whenever $x' \geq x$ and $x' \neq x$.

We proceed with the definition of the parallelogram property, which is the central concept in this paper, followed by the two main results. Throughout, we assume that T is a poset and that $\Gamma : T \to \mathbb{R}^\ell$ is a correspondence taking nonempty set-values $\Gamma(t)$.

Definition (Parallelogram property)**.** The correspondence Γ satisfies the *parallelogram property* for the set $K \subseteq \{1, 2, \ldots, \ell\}$, if for any $t' \geq_T t$ and $x \in \Gamma(t)$, $x' \in \Gamma(t')$, there is $y \in \operatorname{co}\Gamma(t)$, $y' \in \operatorname{co}\Gamma(t')$ such that $x + x' = y + y'$ and $x'_K \geq y_K$, $y'_K \geq x_K$.[6]

If Γ satisfies the parallelogram property for $K = \{1, 2, \ldots, \ell\}$, we shall simply say that it obeys the parallelogram property. The property is depicted in Figure 1a. Given the vectors $x \in \Gamma(t)$, $x' \in \Gamma(t')$, for some $t' \geq_T t$, it is possible to find points y, y' in the convex hulls of $\Gamma(t)$, $\Gamma(t')$, respectively, such that $x' \geq y$ and $y' \geq x$. In particular, the vector y is not an element of $\Gamma(t)$, but it belongs to its convex hull. Moreover, the four elements x, x', y, y' form a parallelogram, so $x + x' = y + y'$.

In contrast, this is not possible in Figure 1b.[7] Indeed, given x, x', for any y, y' such that $x' \geq y$, $y' \geq x$, and $x + x' = y + y'$, either $y \notin \operatorname{co}\Gamma(t) = \Gamma(t)$ or $y' \notin \operatorname{co}\Gamma(t') = \Gamma(t')$.

[6] For any set A, we denote its convex hull by $\operatorname{co} A$.

[7] This correspondence is given by $\Gamma(t) := \{x \in \mathbb{R}^2_+ : ax_1 + (b - t)x_2 \geq tb\}$, for some $a, b > 0$.

In this case, Γ violates parallelogram property for $K = \{1, 2\}$. To be precise, the property fails for $K = \{2\}$, even though it holds for $K = \{1\}$ (essentially because the boundary of $\Gamma(t)$ becomes steeper with t).

With the definition of the parallelogram property in place, we could state our first main result.

Theorem 1. *Let T be a poset and $\Gamma : T \to \mathbb{R}^\ell$ be a correspondence with compact values. For any $K \subseteq \{1, 2, \ldots, \ell\}$, the following statements are equivalent.*

(i) *The correspondence Γ satisfies the parallelogram property for K.*

(ii) *For any $p \in \mathbb{R}^\ell$, the correspondence $\Phi : T \to \mathbb{R}^\ell$, given by*

$$\Phi(t) \ := \ \text{argmin} \Big\{ p \cdot y : y \in \Gamma(t) \Big\}, \tag{1}$$

satisfies the parallelogram property for K.[8]

(iii) *The function $f : \mathbb{R}^\ell \times T \to \mathbb{R}$, given by $f(p, t) := \min \big\{ p \cdot y : y \in \Gamma(t) \big\}$, has increasing differences in (p_K, t), i.e., for any $p'_K \geq p_K$, $t' \geq_T t$, and p_K,*

$$f\big((p'_K, p_{-K}), t'\big) - f\big((p_K, p_{-K}), t'\big) \ \geq \ f\big((p'_K, p_{-K}), t\big) - f\big((p_K, p_{-K}), t\big). \tag{2}$$

Proof. To show (i) $\Rightarrow$ (ii), take any $p \in \mathbb{R}^\ell$, $t' \geq_T t$, and $x \in \Phi(t)$, $x' \in \Phi(t')$, where the two sets are non-empty. Since $x \in \Gamma(t)$, $x' \in \Gamma(t')$, by the parallelogram property, there is some $y \in \text{co}\,\Gamma(t)$, $y' \in \text{co}\,\Gamma(t')$ such that $x + x' = y + y'$ and $x'_K \geq y_K$, $y'_K \geq x_K$. We claim that $y \in \text{co}\,\Phi(t)$ and $y' \in \text{co}\,\Phi(t')$. Since $y \in \text{co}\,\Gamma(t)$ and $x \in \Phi(t)$, it must be that $p \cdot y \geq p \cdot x$.[9] Similarly, we have $p \cdot y' \geq p \cdot x'$. Thus,

$$p \cdot (y + y') \ \geq \ p \cdot (x + x') \ = \ p \cdot (y + y'),$$

which holds only if $p \cdot y = p \cdot x$ and $p \cdot y' = p \cdot x'$, and so $y \in \text{co}\,\Phi(t)$, $y' \in \text{co}\,\Phi(t')$.

We prove that (ii) implies (iii). It is well known that f is a concave function. In particular, the map from $z \in [p_i, p'_i]$ to $f(z, p_{-i}, t)$ is concave and continuous on the interval $[p_i, p'_i]$. Hence it is absolutely continuous and thus almost everywhere differentiable (see Theorem 25.5 in Rockafellar, 1970), with

$$f\big((p'_i, p_{-i}), t\big) - f\big((p_i, p_{-i}), t\big) \ = \ \int_{p_i}^{p'_i} \frac{\partial f}{\partial p_i}\big((z, p_{-i}), t\big) dz.$$

[8] The set $\Phi(t)$ is nonempty because $\Gamma(t)$ is compact.

[9] This is because $\min \big\{ p \cdot y : y \in \Gamma(t) \big\} = \min \big\{ p \cdot y : y \in \text{co}\,\Gamma(t) \big\}$, for any $p \in \mathbb{R}^\ell$ and $t \in T$.

By Theorem 25.1 in Rockafellar (1970) (Shephard's Lemma), whenever $\partial f/\partial p_i\big((z,p_{-i}),t\big)$ exists it is equal to y_i, for any $y \in \Phi(t)$ (with $p = (z,p_{-i})$). Since Φ obeys the parallelogram property, for any $t' \geq_T t$ and $i \in K$, we have $\partial f/\partial p_i(z,p_{-i},t) \leq \partial f/\partial p_i(z,p_{-i},t')$, for almost all $z \in [p_i,p_i']$. This leads to

$$f\big((p_i',p_{-i}),t\big) - f\big((p_i,p_{-i}),t\big) \;\leq\; f\big((p_i',p_{-i}),t'\big) - f\big((p_i,p_{-i}),t'\big).$$

Thus we have shown that the function f has increasing differences in (p_i,t), for any $i \in K$, which implies that f has increasing differences in (p_K,t).

The proof that (iii) implies (i) is rather long and left to the Appendix. Our argument employs the separating hyperplane theorem to show that if the parallelogram property fails, then f violates increasing differences. □

This theorem states, in particular, that the parallelogram property for K is necessary and sufficient for f to have increasing differences in (p_K,q). The sufficiency part of this claim can be proven directly. Indeed, take any $t' \geq_T t$ and $x \in \Gamma(t)$, $x' \in \Gamma(t')$. By assumption, there is some $y \in \operatorname{co}\Gamma(t)$ and $y' \in \operatorname{co}\Gamma(t')$ satisfying $x + x' = y + y'$, and $x'_K \geq y_K$, $y'_K \geq x_K$. Let $p' = (p'_K, p_{-K})$ and $p = (p_K,p_{-K})$, where $p'_K \geq p_K$. Then $p' \cdot \big[x' - y\big] \geq p \cdot \big[x' - y\big] = p \cdot [y' - x]$, which guarantees that

$$p \cdot x + p' \cdot x' \;\geq\; p \cdot y' + p' \cdot y \;\geq\; f(p,t') + f(p',t).$$

Taking the minimum over the left hand-side yields $f(p,t) + f(p',t') \geq f(p,t') + f(p',t)$, which is precisely the property that f satisfies increasing differences in (p_K,t).

Theorem 1 determines comparative statics of minimization problems with an arbitrary linear objective. In some applications we require comparative statics over a narrower class of strictly increasing objective functions; in those cases, the following result is helpful.

Theorem 2. *Let T be a poset and $\Gamma : T \to \mathbb{R}^\ell$ be a correspondence with closed, upward comprehensive,*[10] *and bounded from below values.*[11] *For any $K \subseteq \{1,2,\ldots,\ell\}$, the following statements are equivalent.*

(i) *The correspondence Γ satisfies the parallelogram property for K.*

(ii) *For any strictly positive $p \in \mathbb{R}^\ell_{++}$, the correspondence $\Phi : T \to \mathbb{R}^\ell$ defined as in (1) satisfies the parallelogram property for K.*[12]

[10] Set $\Gamma(t)$ is *upward comprehensive* if $x \in \Gamma(t)$ and $y \geq x$ implies $y \in \Gamma(t)$, for any $y \in \mathbb{R}^\ell$.

[11] Set $\Gamma(t)$ is *bounded from below* if there is some $y \in \mathbb{R}^\ell$ such that $x \geq y$, for all $x \in \Gamma(t)$.

[12] Our assumptions on $\Gamma(t)$ guarantee that $\Phi(t)$ is nonempty.

(iii) *The function* $f : \mathbb{R}^{\ell}_{++} \times T \to \mathbb{R}$*, given by* $f(p,t) := \min\{p \cdot y : y \in \Gamma(t)\}$*, satisfies increasing differences in* (p_K, t)*, defined as in (2).*

The argument that (i) ⇒ (ii) ⇒ (iii) is the same as the one provided for the proof of Theorem 1. The proof that (iii) implies (i) is in the Appendix.

Remark 2.1 (Maximization problems)**.** There is an analogous version of Theorem 1 applicable to maximization problems where the following are equivalent: (i) the correspondence $\Gamma : (T, \geq_T) \to \mathbb{R}^{\ell}$ satisfies the parallelogram property for K; (ii) the correspondence $\tilde{\Phi}(t) := \operatorname{argmax}\{q \cdot y : y \in \Gamma(t)\}$ satisfies the parallelogram property for K, for any $q \in \mathbb{R}^{\ell}$; (iii) the function $\tilde{f}(q,t) := \max\{q \cdot y : y \in \Gamma(t)\}$ has increasing differences in (q_K, t). This result can be obtained by applying Theorem 1 to the correspondence $\Gamma^* : (T, \geq^*_T) \to \mathbb{R}^{\ell}$, where $\Gamma^*(t) = -\Gamma(t)$ and $t' \geq^*_T t$ if $t \geq_T t'$; note that Γ^* has the parallelogram property for K (with respect to $(T, \geq^*_T)$) if and only if Γ has the parallelogram property for K (with respect to $(T, \geq_T)$).

The same trick yields a version of Theorem 2 for maximization problems for any correspondence Γ that has downward (rather than upward) comprehensive and bounded from above (rather than below) values, with $q \in \mathbb{R}^{\ell}_{++}$.

Remark 2.2 (Generalizations)**.** We have focused our discussion on the case where Γ is mapped into the Euclidean space because that is the context of our applications, but it is not difficult to see that this restriction can be relaxed for some results. Indeed the parallelogram property can be defined for any correspondence Γ mapping T to an ordered vector space $(X, \geq_X)$:[13] For any $t' \geq_T t$ and $x \in \Gamma(t), x' \in \Gamma(t')$ we require that there be $y \in \operatorname{co}\Gamma(t), y' \in \operatorname{co}\Gamma(t')$ such that $x + x' = y + y'$, with $x' \geq_X y$, $y' \geq_X x$. An argument analogous to the one in the proof of Theorem 1 will show that, for any linear functional $\phi : X \to \mathbb{R}$, the correspondence $\Phi(t) := \operatorname{argmin}\{\phi(y) : y \in \Gamma(t)\}$ satisfies parallelogram property. In addition, by applying the argument following the proof of Theorem 1 (see page 8), we can show directly that the the parallelogram property on Γ guarantees that $f(\phi, t) := \min\{\phi(y) : y \in \Gamma(t)\}$ has increasing differences in (ϕ, t), i.e., if $\phi'(y) \geq \phi(y)$, for all $y \geq_X 0$, and $t' \geq_T t$, then $f(\phi', t') - f(\phi, t') \geq f(\phi', t) - f(\phi, t)$.

[13] An *ordered vector space* $(X, \geq_X)$ consists of a linear space X over the real numbers $\mathbb{R}$ and a preorder $\geq_X$ that preserves linear transformations, i.e., such that $x \geq_X y$ implies $x + z \geq_X y + z$ and $\alpha x \geq_X \alpha y$, for any $z \in X$ and $\alpha \in \mathbb{R}_+$. A *preorder* is a reflexive and transitive binary relation.

We now discuss the relationship between our results and the monotone comparative statics results by Topkis (1978), Milgrom and Shannon (1994), and Quah (2007). These results are situated in the context of a *lattice*, which is a poset $(X, \geq_X)$ such that, for any elements $x, x' \in X$, both their meet (the greatest lower bound) $x \wedge x'$ and their join (the least upper bound) $x \vee x'$ belong to X. Moreover, we say that the set $Y \subseteq X$ is a *sublattice* of X, if for any $x, x' \in Y$, the set Y contains $x \wedge x'$ and $x \vee x'$, that are defined as for X. One can easily check that $(\mathbb{R}^\ell, \geq)$ is a lattice, with the meet $x \wedge x'$ and the join $x \vee x'$ given by $(x \wedge x')_i = \min\{x_i, x'_i\}$ and $(x \vee x')_i = \max\{x_i, x'_i\}$, for all $i = 1, 2, \ldots, \ell$.

A correspondence $\Gamma : T \to X$, that maps a poset T to a lattice X, increases in the *strong set order* if, for any $t' \geq_T t$ and $x \in \Gamma(t), x' \in \Gamma(t')$, we have $x \wedge x' \in \Gamma(t)$, $x \vee x' \in \Gamma(t')$. Given such a correspondence Γ, Topkis (1978) showed that the correspondence of optimal points $\Phi(t) := \operatorname{argmin}\big\{\phi(y) : y \in \Gamma(t)\big\}$ is also increasing in the strong set order if the objective function $\phi : X \to \mathbb{R}$ is *submodular*, i.e., satisfies $\phi(x) + \phi(x') \geq \phi(x \wedge x') + \phi(x \vee x')$, for any $x, x' \in X$. Observing that any comparative statics result on Φ must be independent of strictly increasing transformations of the objective function, Milgrom and Shannon (1994) generalize Topkis' result by showing that it suffices for ϕ to satisfy the ordinal counterpart of submodularity, which they call *quasisubmodularity*; this property requires the following: if $\phi(x \wedge x') \geq (>)\, \phi(x)$ then $\phi(x') \geq (>)\, \phi(x \vee x')$, for any $x, x' \in X$.

Quah (2007) observes that for certain economic problems, the strong set order on Γ is an overly strong assumption. He develops a comparative statics result that requires an ordinal condition on the objective function ϕ that is stronger than quasisubmodularity (called $\mathcal{C}$-quasisubmodularity[14]), while relaxing the strong set order requirement on Γ. Specifically, the correspondence Γ is required to be increasing in the *$\mathcal{C}$-flexible set order for $K \subseteq \{1, 2, \ldots, \ell\}$*, which means that, for any $t' \geq_T t$, $x \in \Gamma(t)$ and $x' \in \Gamma(t')$ with $x'_K \ngeq x_K$, there is some $\lambda \in [0, 1]$ such that $\big[\lambda x' + (1 - \lambda)(x \wedge x')\big] \in \Gamma(t)$ and $\big[\lambda x + (1 - \lambda)(x \vee x')\big] \in \Gamma(t')$. Obviously, this order is weaker than the strong set order, which corresponds to the case where $\lambda = 0$. It is shown that if Γ increases in the $\mathcal{C}$-flexible set order for K, then so does Φ, for any $\mathcal{C}$-quasisubmodular function ϕ.

In this paper, we push the approach in Quah (2007) even further, by requiring the

[14] A sufficient condition for this property is that ϕ is submodular and convex.

objective ϕ to be *linear*, in order to obtain the most permissive conditions on Γ needed for monotone comparative statics. This leads to the parallelogram property, which is weaker than the $\mathcal{C}$-flexible set order since $x + x' = y + y'$ with $y' \geq x$ and $x' \geq y$, if we let $y = \lambda x' + (1-\lambda)(x \wedge x')$ and $y' = \lambda x + (1-\lambda)(x \vee x')$.

3 Comparative statics of firms and consumers

In this section we first characterize those technologies which guarantee that a firm has normal demand for inputs and marginal cost that increases with input prices. We then go on to examine the closely related issue of normal demand for consumers.

3.1 Normal input demand and increasing marginal costs

We assume the firm's output is governed by the production function $F : X \to \mathbb{R}$; to keep the exposition short, we shall assume throughout this section that $X = \mathbb{R}^\ell_+$ and that F is continuous. We denote the range of F by $Q := \left\{F(x) : x \in \mathbb{R}^\ell_+\right\}$. For any $q \in Q$, the set $U(q) := \left\{x \in \mathbb{R}^\ell_+ : F(x) \geq q\right\}$; clearly, this set is bounded from below and closed.

We assume that the firm is a price-taker in the market for inputs, facing strictly positive input prices $p = (p_1, p_2, \ldots, p_\ell)$ in $\mathbb{R}^\ell_{++}$. The *input* or *factor demand* at p and output $q \in Q$ refers to those bundles that achieve output of at least q with the least cost. Formally, input demand is the correspondence $H : \mathbb{R}^\ell_{++} \times Q \to \mathbb{R}^\ell_+$, where

$$H(p,q) \ := \ \text{argmin}\left\{p \cdot x : F(x) \geq q\right\}.$$

Our assumptions on F guarantee that $H(p,q)$ is nonempty and compact for all (p,q) in $\mathbb{R}^\ell_{++} \times Q$. The associated cost function is $C : \mathbb{R}^\ell_{++} \times Q \to \mathbb{R}_+$, where

$$C(p,q) \ := \ \min\left\{p \cdot x : F(x) \geq q\right\}.$$

Obviously, we have $C(p,q) = p \cdot x$ for any $x \in H(p,q)$.

We are interested in the conditions under which an increase in the price of inputs i in $K \subseteq \{1, 2, \ldots, \ell\}$ leads to higher marginal cost. In the case where C is not necessarily differentiable, this means that for any prices $p'_K \geq p_K$, p_{-K}, and output levels $q' \geq q$,

$$C\big((p'_K, p_{-K}), q'\big) - C\big((p'_K, p_{-K}), q\big) \ \geq \ C\big((p_K, p_{-K}), q'\big) - C\big((p_K, p_{-K}), q\big). \qquad (3)$$

Thus, the increase in cost when output is raised from q to q' is greater at the input prices (p'_K, p_{-K}) compared to the prices (p_K, p_{-K}) when $p'_K \geq p_K$. When C is differentiable in output, it is straightforward to show that C has increasing differences in (p_K, q) if and only if $\partial C/\partial q\big((p_K, p_{-K}), q\big)$ increases in p_K. Note also that the increasing differences property is additive across inputs; i.e., C has increasing differences in (p_K, q) if and only if it has increasing differences in (p_i, q), for all $i \in K$.

Increasing marginal costs are closely related to normality of the input demand, defined as follows: the input demand correspondence H is *normal* in $K \subseteq \{1, 2, \ldots, \ell\}$ if for any input prices p, output levels q, q', and input profile $x \in H(p, q)$, there is some input profile $x' \in H(p, q')$ such that $q' \geq q$ implies $x'_K \geq x_K$ and $q' \leq q$ implies $x'_K \leq x_K$.

The correspondence H is normal in input i if it is normal in $K = \{i\}$. Moreover, for $K = \{1, 2, \ldots, \ell\}$, we abbreviate by simply saying that H is normal. It follows immediately that H is normal in K only if it is normal in $K' \subseteq K$. In particular, it is normal in input i, for all $i \in K$. When H is a *function*, the converse is also true, but not if it is a correspondence, as we shall show with an example later in this section.

In the following result we characterize both normal input demand and increasing marginal costs. First, let $\overline{U}(q) := \big\{y \in \mathbb{R}^\ell : y \geq x, \text{ for some } x \in U(q)\big\}$ denote the upward comprehensive hull of the super-level set $U(q)$. We say that the function F satisfies the *parallelogram property* for $K \subseteq \{1, 2, \ldots, \ell\}$, if for any $x, x' \in X$ such that $F(x) \geq q$, $F(x') \geq q'$, there is some $y \in \operatorname{co} \overline{U}(q)$, $y' \in \operatorname{co} \overline{U}(q')$ satisfying $x + x' = y + y'$ and $x'_K \geq y_K$, $y'_K \geq x_K$. It is straightforward to check that this equivalent to the correspondence mapping q to $\overline{U}(q)$ having the parallelogram property for K. In the case where $K = \{1, 2, \ldots, \ell\}$, we simply say that F obeys the parallelogram property.

Proposition 1 (Normal input demand)**.** *For an arbitrary continuous production function $F : \mathbb{R}^\ell_+ \to \mathbb{R}$ and subset $K \subseteq \{1, 2, \ldots, \ell\}$, the following statements are equivalent.*

(i) *F satisfies the parallelogram property for K.*

(ii) *For each $p \in \mathbb{R}^\ell_{++}$, the correspondence mapping $q \in Q$ to $H(p, q)$ satisfies the parallelogram property for K.*

(iii) *H is normal in input i, for each $i \in K$.*

(iv) *The cost function C has increasing differences in (p_K, q), as in (3).*

Remark 3.1. This proposition provides a general equivalence between the parallelogram property and marginal costs that increase with input prices. On the other hand, the relationship between the parallelogram property and input demand is more subtle. While the parallelogram property for K guarantees that the correspondence H is normal for each input $i \in K$, it does *not* guarantee that H is normal in K. For example, consider the production function F with isoquants depicted in Figure 2. Clearly, it obeys the parallelogram property, essentially because $\operatorname{co} U(q) = \{(x_1, x_2) : x_1 + x_2 \geq q\}$, for all $q \in Q$. In line with the proposition, the input demand correspondence is normal in input 1 and in input 2 (separately). Specifically, for any input prices such that $p_1 = p_2$ and given $x \in H(p, q)$, there is $x' \in H(p, q')$ such that $x'_1 \geq x_1$ and $x'' \in H(p, q')$ such that $x''_2 \geq x_2$. However, F is not quasiconcave, the super-level set $U(q') = \{x \in \mathbb{R}^\ell_+ : F(x) \geq q'\}$ is not convex and there is no $y \in H(p, q')$ such that $y \geq x$.

Remark 3.2. Suppose that the production function F is quasiconcave. The proposition states that if F satisfies the parallelogram property for K then $H(p, \cdot)$ also satisfies the parallelogram property for K, for all p. This in turn guarantees that H *is normal in* K (rather than just normal for each $\{i\}$ in K) since $\operatorname{co} H(p, q) = H(p, q)$. Furthermore, we show in the Appendix that the parallelogram property on H would then guarantees that H has a selection that is normal in K, i.e., there is a function $h : \mathbb{R}^\ell_{++} \times Q \to \mathbb{R}^\ell$ such that $h(p, q) \in H(p, q)$, for all (p, q) in $\mathbb{R}^\ell_{++} \times Q$, and $q' \geq q$ implies $h_K(p, q') \geq h_K(p, q)$, for any $p \in \mathbb{R}^\ell_{++}$.

Proof of Proposition 1. To prove that (i) $\Rightarrow$ (ii), we first note that if F satisfies the parallelogram property for K, then so does the correspondence $q \to \overline{U}(q)$. For any $p \in \mathbb{R}^\ell_{++}$, Theorem 2 guarantees that the correspondence mapping $q \in Q$ to $H(p, q)$ also satisfies the parallelogram property for K. To see that (ii) $\Rightarrow$ (iii), note that if $H(p, \cdot)$ satisfies the parallelogram property for K, then for any $q' \geq q$ and $x \in H(p, q)$ there is $y' \in \operatorname{co} H(p, q')$ such that $y'_K \geq x_K$. Thus there are vectors $z^j \in H(p, q')$ and numbers $\alpha^j \geq 0$, for $j = 1, \ldots, m$, such that $y' = \sum_{j=1}^m \alpha^j z^j$ and $\sum_{j=1}^m \alpha^j = 1$. Since $y'_K \geq x_K$, there is some j with $z^j_i \geq x_i$. An analogous argument establishes that, for any $x' \in H(p, q')$ and $i \in K$, there is $z \in H(p, q)$ such that $x'_i \geq z_i$. The implication (iii) $\Rightarrow$ (iv) can be shown as in the proof of Theorem 1, by applying Shephard's lemma. Finally, the implication (iv) $\Rightarrow$ (i) follows from Theorem 2. $\square$

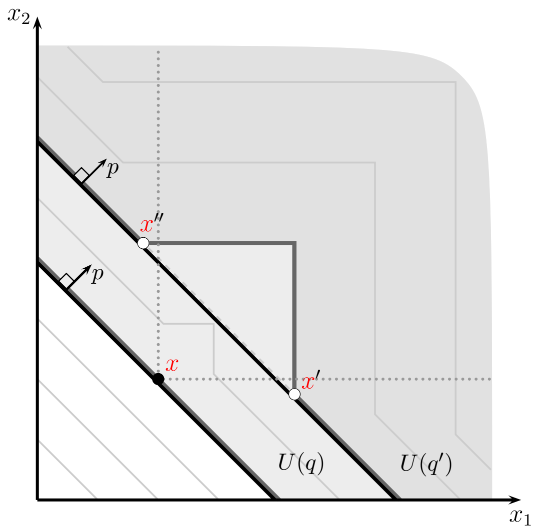


Figure 2: Normal input demand.

Proposition 1 states that when F has the parallelogram property for i, an increase in the price of i raises marginal cost. Our intuition suggests that this should lower the firm's output. To state this precisely, we require F to be increasing (in addition to being continuous). Then for any $x \in H(p,q)$, we have $F(x) = q$ and the firm's profit-maximizing choice could be conveniently broken down into two steps: first it figures out $C(p,q)$ and then it chooses $q \in Q$ to maximize $R(q) - C(p,q)$, where $R : Q \to \mathbb{R}$ is the revenue attracted by q. Formally, if we let $\bar{Q}(p) := \text{argmax}_{q \in Q} \{R(q) - C(p,q)\}$ and $\bar{X}(p) := \text{argmax}_{x \in \mathbb{R}^\ell_+} \{R(F(x)) - p \cdot x\}$, then the following holds:

(A) if $\hat{x} \in \bar{X}(p)$, then $F(\hat{x}) \in \bar{Q}(p)$ and $\hat{x} \in H(p, F(\hat{x}))$, and

(B) if $\hat{q} \in \bar{Q}(p)$ and $\hat{x} \in H(p,\hat{q})$, then $\hat{x} \in \bar{X}(p)$ and $F(\hat{x}) = \hat{q}$.

How does the firm's optimal output, i.e., the set $\{F(x') : x' \in \bar{X}(p)\}$, vary with p_i? Observe that (A) and (B) together guarantee that this set coincides with $\bar{Q}(p)$. Furthermore, if F has the parallelogram property for i then $R(q) - C(p,q)$ has decreasing differences in (p_i, q); this in turn guarantees that $\bar{Q}(p)$ decreases with p_i in the strong set order (see Topkis, 1978 or Milgrom and Shannon, 1994), thus confirming our intuition.

Another implication of Proposition 1 is that when F has the parallelogram property for i then the conditional factor demand for i is normal. This suggests that any change in market conditions that stimulate output (such as an increase in the price of the output in the case where the firm is a price-taker) will raise the demand for factor i.

In summary, the following threefold equivalence is a consequence of Proposition 1.

Proposition 2. *For any continuous and increasing production function $F : X \to \mathbb{R}$, the following statements are equivalent.*

(i) *F satisfies the parallelogram property for i.*

(ii) *The set $\{F(x) : x \in \text{argmax}_{x' \in X}\{R(F(x')) - p \cdot x'\}\}$ is decreasing in $p_i > 0$ in the strong set order, for any increasing function $R : Q \to \mathbb{R}$.*

(iii) *Let T be a poset and $R(\cdot, t) : Q \to \mathbb{R}$ be a family of increasing functions with increasing differences in (q, t). The correspondence $\widehat{X}(t) : T \to X$, given by*

$$\widehat{X}(t) \;:=\; \text{argmax}\Big\{R\big(F(x), t\big) - p \cdot x : x \in X\Big\},$$

increases in the following sense: for any $t', t'' \in T$ and $\hat{x}' \in \widehat{X}(t')$ there is some $\hat{x}'' \in \widehat{X}(t'')$ such that $t'' \geq_T t'$ implies $\hat{x}''_i \geq \hat{x}'_i$ and $t'' \leq_T t'$ implies $\hat{x}''_i \leq \hat{x}'_i$.

The proof of this result is in the Appendix. In statement (iii), the property that $R(\cdot, t)$ has increasing differences in (q, t) guarantees that the firm's profit-maximizing output levels, which is equal to $\text{argmax}_{q \in Q}\{R(q,t) - C(p,q)\}$, increases with t. The equivalence of (ii) and (iii) in Proposition 2 tells us something that is, a priori, unobvious: the complementary relationship between a factor i and output is *symmetric*, in the sense that a fall in the price of factor i always stimulates output if and only if any change in market condition that stimulates output also raises the usage of i.[15]

3.2 Related results

Some version of the equivalence of statements (iii) and (iv) in Proposition 1 is known, at least since Fisher (1990). Fisher's original argument assumes differentiability of the objective function F and the global uniqueness of demand; we dispense with these assumptions altogether. Our argument that (iii) implies (iv) does not significantly break new ground, but it is worth emphasizing that our proof that (iv) implies (iii) goes through statement (i) and thus hinges on our characterization of normality using the parallelogram property. The parallelogram property is a geometrically intuitive generalization

[15] In the classical theory of a price-taking firm, if the firm's profit function is C^2, then the optimal choice of the firm is unique and differentiable; denoting the optimal output by $\bar{q}(p, m)$ (where m is the output price) and the optimal demand for i by $\bar{x}_i(p, m)$, one obtains (via Hotelling's Lemma and the symmetry of second derivatives) that $d\bar{q}/dp_i = d\bar{x}_i/dm$ and, in particular, they have the same sign. The equivalence between (ii) and (iii) is a manifestation of this phenomenon in a more general setting.

of known conditions that lead to normal demand, such as homotheticity and the case of supermodular and concave utility/production functions (see Chipman, 1977 or Quah, 2007); we discuss this connection in greater detail in the next subsection.

Alarie et al. (1990) and Bilancini and Boncinelli (2010) also characterize normal demand, under the condition that the objective function $F : \mathbb{R}^{\ell}_{++} \to \mathbb{R}$ is strictly increasing, strictly quasiconcave, and twice-differentiable. Note that these conditions on F are sufficient for demand to be strictly positive for all goods and a smooth *function* of prices. Let $G(x)$ and $H(x)$ denote the gradient and the Hessian of F, respectively, at some bundle x. The corresponding bordered Hessian $\tilde{H}(x)$ is given by

$$\tilde{H}(x) \;:=\; \begin{bmatrix} H(x) & G(x) \\ G'(x) & 0 \end{bmatrix},$$

where $G'(x)$ is the transposition of the column vector $G(x)$. Let $\tilde{H}_{i,j}(x)$ denote the (i,j)'th minor of $\tilde{H}(x)$ and $\left|\tilde{H}_{i,j}(x)\right|$ be its determinant. Bilancini and Boncinelli (2010) show that the demand function is normal in $K \subseteq \{1, 2, \ldots, \ell\}$ if and only if

$$(-1)^{i-1}\left|\tilde{H}_{i,(\ell+1)}(x)\right| \;\geq\; 0, \tag{4}$$

for all $i \in K$ and each $x \in \mathbb{R}^{\ell}_{++}$. By combining Proposition 1 with this result, we conclude that the condition (4) is equivalent to the parallelogram property for K, when F satisfies the ancillary smoothness assumptions in their setup.[16]

3.3 Technologies that generate normal factor demand

In this section we give a quick survey of some production functions that lead to normal demand and relate them to our results.

Example 1. It is well-known that if the production function F is homothetic then its input demand is linear in output, which implies normality. Thus homothetic production functions must obey the parallelogram property (by Proposition 1); we provide a direct proof of this claim in the Online supplement.

[16] In the case where there are just two goods, there is a characterization of normality provided by Cherchye et al. (2018) in a revealed preference framework à la Afriat (1967). Given a finite set of expenditure data, they find necessary and sufficient conditions under which the data set can be rationalized by a utility function that generates normal demand.

Example 2. Let X be a convex sublattice of $\mathbb{R}^\ell$. We say that a function $F : X \to \mathbb{R}$ is increasing in the $\mathcal{C}$-flexible set order for $K \subseteq \{1, 2, \dots, \ell\}$ if the correspondence $q \mapsto U(q)$ is increasing in the $\mathcal{C}$-flexible set order for K; this property implies that F has the parallelogram property.[17] It is known that F is increasing in the $\mathcal{C}$-flexible set order for K if it is continuous, increasing, supermodular, and concave in x_{-i}, for all $i \in K$.[18] Note that not every function that has the parallelogram property is increasing in the $\mathcal{C}$-flexible set order; for example, $F(x_1, x_2, x_3) := \min\{x_1, x_2 + x_3\}$ is homogenous of degree 1 and thus obeys the parallelogram property, but it is not increasing in the $\mathcal{C}$-flexible set order.[19]

Suppose $\ell = 2$ with $X := X_1 \times X_2$, where X_1, X_2 are intervals of $\mathbb{R}$; a continuously differentiable and strictly increasing function $F : X \to \mathbb{R}$ is increasing in the $\mathcal{C}$-flexible set order for $\{i\}$ (and thus satisfies the parallelogram property for $\{i\}$) if

$$-\left[\frac{\partial F}{\partial x_i}(x_1, x_2) \bigg/ \frac{\partial F}{\partial x_j}(x_1, x_2)\right] \tag{5}$$

(the marginal rate of substitution) is decreasing in x_j, for $j \neq i$.[20] This is illustrated in Figure 1b, where the indifference curves become steeper as x_2 increases, for fixed x_1.

Example 3. There is a way of constructing new production functions that generate normal demand from other production functions with that property. Let $f^j : \mathbb{R}^\ell_+ \to \mathbb{R}$ (for $j = 1, 2 \dots, n$) be a collection of functions and define $F : \mathbb{R}^\ell_+ \to \mathbb{R}$ by

$$F(x) \;:=\; \max\left\{G\Big(f^1(y^1), f^2(y^2), \dots, f^n(y^n)\Big) : x \geq \sum_{j=1}^n y^j\right\}, \tag{6}$$

where $G : \times_{j=1}^n Q^j \to \mathbb{R}$ is some function that aggregates the values of f^j and Q^j contains the range of f^j. For example, if $G(q^1, q^2, \dots, q^n) = \sum_{k=1}^n q^k$, then F can be interpreted as the production function of a firm that allocates its output efficiently across n plants, with plant j having the production function f^j. In the Online supplement we show that

[17] Recall the discussion on the $\mathcal{C}$-flexible set order on page 10 at the end of Section 2.

[18] For a proof of this claim see Quah (2007) or the Online supplement. For example, consider $F(x_1, x_2, x_3) := \sqrt{x_1 x_2 x_3}$ for $(x_1, x_2, x_3) \in \mathbb{R}^3_+$. This function is continuous, increasing, supermodular, and concave in each pair of variables ((x_1, x_2), (x_1, x_3), and (x_2, x_3)), though it is not concave in the three variables jointly. Thus it is increasing in the $\mathcal{C}$-flexible set order for $K = \{1, 2, 3\}$.

[19] If $x = (5, 0, 5)$ and $x' = (6, 6, 0)$, then $F(x) = 5$, $F(x') = 6$, $F(x \wedge x') = F\big((5, 0, 0)\big) = 0$, and $F(x \vee x') = F\big((6, 6, 5)\big) = 6$. For F to be increasing in the $\mathcal{C}$-flexible set order, we must find $\lambda \in [0, 1]$ such that $F\big(\lambda x + (1 - \lambda)(x \wedge x')\big) \geq F(x) = 5$ and $F\big(\lambda x + (1 - \lambda)(x \vee x')\big) \geq F(x') = 6$. But this is impossible since $F(x \wedge x') < 5$ and if $\lambda > 0$, $F\big(\lambda x + (1 - \lambda)(x \vee x')\big) < 6$.

[20] This is based on the Supplemental Appendix in Quah (2007); see Online supplement for the details.

F has the parallelogram property for K provided f^j is continuous, concave, and has the parallelogram property for K (for every j), and the aggregating function G is increasing in the $\mathcal{C}$-flexible set order. We also provide examples to show that F need not be homothetic or increasing in the $\mathcal{C}$-flexible set order, even when these properties hold for all f^j.

3.4 Consumer demand

Our results on conditional factor demand can be straightforwardly re-formulated to guarantee that Marshallian demand is normal. Apart from being an intrinsically appealing property for most product categories, normality plays an important role in many model settings. For example, normality is used in Bergstrom et al. (1986) to guarantee the uniqueness of Nash equilibria in a public goods game; the results on general equilibrium comparative statics in Nachbar (2002) and Quah (2003) hinge on their assumption that demand is normal; in Blundell et al. (2005), normality helps to determine how provision of a public good varies with intra-household bargaining power; and normality simplifies the non-parametric estimation of demand functions in Blundell et al. (2003). Thus it is important to have a through understanding of the foundations of this property.

Suppose a consumer has a utility function $u : \mathbb{R}^\ell_+ \to \mathbb{R}$, defined over bundles of ℓ commodities. At prices $p \in \mathbb{R}^\ell_{++}$ and income $m \geq 0$, the consumer chooses a consumption bundle $x \in \mathbb{R}^\ell_+$ that is affordable and maximizes her utility; the solution to this problem is captured by the *Marshallian demand correspondence* $D : \mathbb{R}^\ell_{++} \times \mathbb{R}_+ \to \mathbb{R}^\ell_+$, where $D(p, m) := \text{argmax} \left\{u(x) : p \cdot x \leq m\right\}$. The indirect utility function $v : \mathbb{R}^\ell_{++} \times \mathbb{R}_+ \to \mathbb{R}$ is given by $v(p, m) := \max \left\{u(x) : p \cdot x \leq m\right\}$.

Given the range V of u, the *Hicksian demand* $H : \mathbb{R}^\ell_{++} \times V \to \mathbb{R}^\ell_+$ maps prices p and utility levels v to those bundles that minimize the expenditure $p \cdot x$ over all alternatives satisfying $u(x) \geq v$. Obviously, the Hicksian demand is formally identical to the input demand in the production context, while the analog to the cost function is the expenditure function $e : \mathbb{R}^\ell_{++} \times V \to \mathbb{R}_+$, where $e(p, v) := \min \left\{p \cdot x : u(x) \geq v\right\}$.

Suppose that utility u is continuous and locally non-satiated.[21] In such a case, correspondences D and H are well-defined. Moreover, we have $p \cdot x = m$, for all $x \in D(p, m)$,

[21] Utility function u is *locally non-satiated* if, for any bundle $x \in \mathbb{R}^\ell_+$, there is another bundle y arbitrarily close to x such that $u(y) > u(x)$.

while the two demands are related by the identity $D(p, m) = H\big(p, v(p, m)\big)$, for any prices p and income m (see Proposition 3.E.1 in Mas-Colell et al., 1995).

Let $K \subseteq \{1, 2, \ldots, \ell\}$. We say that D is *normal* in K if, for any prices p, income levels m, m', and $x \in D(p, m)$, there is $x' \in D(p, m')$ such that $m' \geq m$ implies $x'_K \geq x_K$ and $m' \leq m$ implies $x'_K \leq x_K$. The demand D is normal in input i if it is normal in $K = \{i\}$. Finally, if $K = \{1, 2, \ldots, \ell\}$, we simply say that D is normal.

The equivalence of demands D and H allows us to translate normality results on Hicksian demand into results on Marshallian demand. In particular, the following result follows immediately from Proposition 1 and Remark 3.2.

Proposition 3. *Let utility $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ be continuous, locally non-satiated, and quasi-concave. For any set $K \subseteq \{1, 2, \ldots, \ell\}$ the following statements are equivalent.*

(i) *The function u satisfies the parallelogram property for K.*

(ii) *For any prices $p \in \mathbb{R}^\ell_{++}$, the correspondence mapping $m \in \mathbb{R}_+$ to $D(p, m)$ satisfies the parallelogram property for K.*

(iii) *D is normal in K.*[22]

It is well-known that if a Marshallian demand function d is normal for good i, then the demand for i obeys the *law of demand*, i.e., function $d_i\big((p_i, p_{-i}), m\big)$ is decreasing in p_i, for all p_{-i} and m. We know that if, for all $p \in \mathbb{R}^\ell_{++}$, the Marshallian demand $D(p, \cdot)$ has the parallelogram property for K, then D admits a selection d that satisfies normality (see Remark 3.2). Therefore, an immediate consequence of Proposition 3 is the following: for any continuous, locally non-satiated, and quasiconcave utility $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ that satisfies the parallelogram property for K, there is a selection $d(p, m)$ from the $D(p, m)$ that obeys the law of demand for every good $i \in K$.

Example 4. Consider an agent who lives for ℓ periods and has a preference over consumption streams $(x_1, x_2, \ldots, x_\ell) \in \mathbb{R}^\ell_+$. In this context, it is natural to assume that this agent's utility has a recursive form, where

$$u(x) \;:=\; h_1\Bigg(x_1, h_2\bigg(x_2, h_3\Big(x_3, \ldots h_{\ell-1}\big(x_{\ell-1}, x_\ell\big)\Big)\bigg)\Bigg)$$

[22] If u is not assumed to be quasiconcave, then (i), (ii) and (iii) are still equivalent if (iii) is modified as follows: D is normal in i, for each $i \in K$. This is clear from Proposition 1.

We claim that when $h_i : \mathbb{R}^2_+ \to \mathbb{R}_+$ are continuous, increasing, concave, and supermodular, then u obeys the parallelogram property, which in turn guarantees that Marshallian demand is normal. Indeed, the properties imposed on h_i guarantee that h_i is increasing in the $\mathcal{C}$-flexible set order; consequently, the map from $(x_{\ell-2}, x_{\ell-1}, x_\ell)$ to $h_{\ell-2}\big(x_{\ell-2}, h_{\ell-1}(x_{\ell-1}, x_\ell)\big)$ has the parallelogram property[23] and, furthermore, this map is concave because both $h_{\ell-2}$ and $h_{\ell-1}$ are concave functions. Repeating this argument, we eventually conclude that u is concave and has the parallelogram property.

4 First order stochastic dominance under ambiguity

In this section, we consider an agent making decisions in an uncertain environment. Suppose that the possible states of the world are represented by a set $S \subseteq \mathbb{R}$; to keep our exposition focused on the essentials we assume that the set $S = \{s_1, s_2, \ldots, s_\ell, s_{\ell+1}\}$ is finite, where $s_1 < s_2 < \ldots < s_\ell < s_{\ell+1}$. We denote the distributions on S by $\triangle_S$. A distribution λ *first order stochastically dominates* distribution μ if $\lambda(s) \leq \mu(s)$ for all $s \in S$; we denote this by $\lambda \succeq \mu$. An important feature of the space of distributions $\triangle_S$ ordered with respect to the first order stochastic dominance is that it is a lattice. In particular, for any distributions λ, λ' their meet and join are defined by $(\lambda \wedge \lambda')(s) = \max\big\{\lambda(s), \lambda'(s)\big\}$ and $(\lambda \vee \lambda')(s) = \min\big\{\lambda(s), \lambda'(s)\big\}$, for all $s \in S$, respectively.[24]

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

[23] We obtain this by applying the result in Example 3 with $G = h_{\ell-2}$, $f^1(x_{\ell-2}, x_{\ell-1}, x_\ell) = x_{\ell-2}$, and $f^2(x_{\ell-2}, x_{\ell-1}, x_\ell) = h_{\ell-1}(x_{\ell-1}, x_\ell)$.

[24] For $(\triangle_S, \succeq)$ to be a lattice, it is crucial that S is a subset of $\mathbb{R}$. Although first order stochastic dominance can be naturally extended to distributions over multi-dimensions spaces, in such a case the pair $(\triangle_S, \succeq)$ would no longer constitute a lattice.

$g(x,s)$ has increasing differences in (x,s) (equivalently, is supermodular in (x,s)) and λ is ordered by first order stochastic dominance in the sense that $\lambda(\cdot,t') \succeq \lambda(\cdot,t)$ whenever $t' \geq t$. In such a case, $x' \geq x$ implies

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

will be increasing in t since $s \to \big[g(x',s) - g(x,s)\big]$ is increasing in s. In other words, f has increasing differences in (x,t), which guarantees that $\operatorname{argmax}\big\{f(x,t) : x \in X\big\}$ increases with t in the strong set order (see Topkis, 1978 or Milgrom and Shannon, 1994).

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

4.1 First order stochastic dominance in the maxmin model

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

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

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

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

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

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

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

This proposition gives us one natural way of defining first order stochastic dominance between sets of distributions since (like the original definition) it characterizes increasing shifts in utility when u is increasing. A natural follow-up question is whether property (F1) is also sufficient to guarantee monotone comparative statics. More precisely, suppose the utility from choosing action x is $g(x, s)$ whenever state s is realized and that g is supermodular. Then one could ask if (F1) guarantees that

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

has increasing differences in (x, t).

It is clear that (F1) would indeed be sufficient in certain special cases of g. For example, suppose X consists of only two actions – 0 and 1 – with $g(1, s)$ increasing in s and $g(0, s)$ decreasing in s, then obviously $f(1, t) - f(0, t)$ is increasing in t if Λ satisfies (F1), since $f(1, t)$ and $f(0, t)$ are separately increasing and decreasing in t. In the study of global games with ambiguity by Ui (2015), this is precisely the assumption imposed on (what we call) g, which then allows the author to conclude that the higher action is chosen by players in the game when they receive a higher signal.

In more general modelling contexts, where there are multiple actions or where $g(1, s)$ and $g(0, s)$ do not vary with s in opposite directions, (F1) is *not* sufficient to guarantee that $f(x, t)$ has increasing differences in t, as we show in the following example.

Example 5. Suppose that $X = \{0, 1\}$ and $S = \{s_1, s_2, s_3\}$. The distribution λ is given by $\lambda(s_1) = 1/2$ and $\lambda(s_2) = 3/4$. The distribution λ' is given by $\lambda'(s_1) = \lambda'(s_2) = 1/2$ and $\mu(s_1) = 1/4$, $\mu(s_2) = 7/8$. Suppose that $T = \{t, t'\}$, where $t' > t$, and $\Lambda(t') = \{\lambda'\}$ and $\Lambda(t) = \text{co}\{\lambda, \mu\}$ is the convex hull of λ' and μ. Since $\lambda' \succeq \lambda$, correspondence Λ obeys stochastic dominance in the sense given by (F1). Let $g : X \times S \to \mathbb{R}$ be such that $g(0, s_1) = g(0, s_2) = 5$, $g(0, s_3) = 21$, $g(1, s_1) = 0$, $g(1, s_2) = 8$, and $g(1, s_3) = 24$; note that $g(x, s)$ is increasing in s and supermodular in (x, s). Since $\int_S g(0, s)d\lambda'(s) > \int_S g(1, s)d\lambda'(s)$, we have $\{0\} = \text{argmax}\{f(x, t') : x \in X\}$. Furthermore, since g is supermodular, we obtain $\int_S g(0, s)d\lambda(s) > \int_S g(1, s)d\lambda(s)$; however, this does not mean that the agent chooses action 0 at $\Lambda(t)$; In fact, since

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

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

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

Proposition 5. *Suppose that correspondence $\Lambda : T \to \triangle_S$ has compact and convex values. The following statements are equivalent.*

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

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

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

(ii) *The function $f : X \times T \to \mathbb{R}$, given by (7), has increasing differences in (x, t), for any supermodular function $g : X \times S \to \mathbb{R}$.*

Remark 4.1. For a general objective function $f(x,t) : X \times T \to \mathbb{R}$ (where $X \subset \mathbb{R}$), a well-known ordinal condition which is sufficient and (in some sense) necessary to guarantee that $\text{argmax}_{x \in X} f(x,t)$ increases in the strong set order with t is that f has single crossing differences (see Milgrom and Shannon, 1994); this property (which is obviously implied by increasing differences) requires that if $f(x',t) \geq (>) f(x,t)$ then $f(x',t') \geq (>) f(x,t')$, for any $x' \geq x$ and $t' \geq_T t$. We show in the Appendix that statements (i) and (ii) in Proposition 5 are also equivalent to the following statement: *The function $f : X \times T \to \mathbb{R}$, given by (7), has single-crossing differences in (x,t), for any supermodular function $g : X \times S \to \mathbb{R}$.* In other words, it is not possible to obtain a condition weaker than (F2) even if all we are interested in is guaranteeing that f satisfies single-crossing differences.

Remark 4.2. We show in the Appendix that Proposition 5 remains true if S is a compact interval of $\mathbb{R}$ and function $g(x,\cdot)$ is Riemann-Stieltjes integrable with respect to each $\lambda \in \Lambda(t)$, for all $x \in X$ and $t \in T$. This holds if either of the following conditions is satisfied: (a) function $g(x,s)$ is continuous in $s \in S$; (b) $g(x,s)$ is bounded on S and has only finitely many discontinuities in s, and all distributions in $\Lambda(t)$ are atomless; or (c) $g(x,s)$ is bounded on S and monotone, and all distributions in $\Lambda(t)$ are atomless.

Remark 4.3. Proposition 5 can be equivalently formulated as saying that property (F2) is necessary and sufficient for the function

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

to have increasing differences in (x,t), for any supermodular function g.[25]

The α-maxmin model by Ghirardato et al. (2004) allows for both ambiguity averse and ambiguity loving behavior, with the agent's utility function having the form

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

for some $\alpha \in [0,1]$. This function has increasing differences in (x,t) if Λ satisfies (F2), since both elements of the sum have increasing differences; of course this in turn guarantees that the set of optimal actions increases with t in the strong set order.

Proof of Proposition 5. To prove (i) $\Rightarrow$ (ii), define correspondence $\Gamma : T \to \mathbb{R}$ by

$$\Gamma(t) \ := \ \Big\{y \in \mathbb{R}^\ell : y_i = -\lambda(s_i), \text{ for all } i = 1,2,\dots,\ell \text{ and } \lambda \in \Lambda(t)\Big\}.$$

It is straightforward to check that Γ satisfies parallelogram property if and only if Λ satisfies (F2). For any function $g : X \times S \to \mathbb{R}$ and distribution λ,

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

Given $x' \geq x$, we define $p, p' \in \mathbb{R}^\ell$ by $p_i = g(x,s_{i+1}) - g(x,s_i)$ and $p'_i = g(x,s_{i+1}) - g(x,s_i)$, for $i = 1,2,\dots,\ell$. Then (8) gives

$$f(x,t') - f(x,t) \ = \ \min\Big\{p\cdot y : y \in \Gamma(t')\Big\} - \min\Big\{p\cdot y : y \in \Gamma(t)\Big\}, \qquad (9)$$

with a similar formula for $f(x',t') - f(x',t)$. If g is supermodular, then $p' \geq p$ and Theorem 1 guarantees that

$$\begin{aligned}&\min\Big\{p\cdot y : y \in \Gamma(t')\Big\} - \min\Big\{p\cdot y : y \in \Gamma(t)\Big\} \\ &\qquad\qquad \leq \ \min\Big\{p'\cdot y : y \in \Gamma(t')\Big\} - \min\Big\{p'\cdot y : y \in \Gamma(t)\Big\}.\end{aligned}$$

Thus, $f(x,t') - f(x,t) \leq f(x',t') - f(x',t)$, and so f has increasing differences.

[25] Indeed, Proposition 5 guarantees that $\min\{\int_S -g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\}$ has decreasing differences in (x,t) since $-g(x,s)$ is submodular; therefore $\max\{\int_S g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\} = -\min\{\int_S -g(x,s)d\lambda(s) : \lambda \in \Lambda(t)\}$ has increasing differences in (x,t).

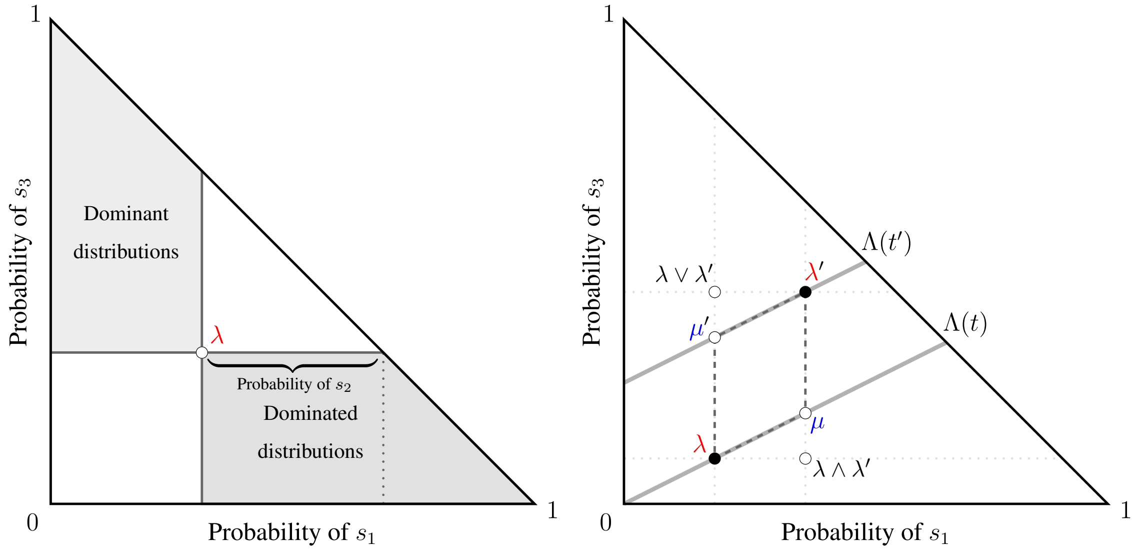


Figure 3: Probability measures represented in the Machina-Marshak triangle. On the right, the thick straight lines represent values $\Lambda(t)$ and $\Lambda(t')$ from Example 7.

We now show that (ii) $\Rightarrow$ (i). If Λ violates (F2) then Γ violates the parallelogram property. By Theorem 1, there are vectors $p' \geq p$ in $\mathbb{R}^\ell$ and $t' \geq_T t$ such that

$$\min\left\{p \cdot y : y \in \Gamma(t')\right\} - \min\left\{p \cdot y : y \in \Gamma(t)\right\}$$
$$> \min\left\{p' \cdot y : y \in \Gamma(t')\right\} - \min\left\{p' \cdot y : y \in \Gamma(t)\right\}. \quad (10)$$

Take any $x, x' \in X$ satisfying $x' > x$. Define a supermodular function $g : X \times S \to \mathbb{R}$ as follows: $g(y, s_1) = 0$ for all $y \in X$ and

$$g(y, s_i) := \begin{cases} \sum_{j=1}^{i-1} p_i & \text{if } y \leq x; \\ \sum_{j=1}^{i-1} p'_i & \text{otherwise,} \end{cases} \quad (11)$$

for $i = 2, 3, \ldots, (\ell + 1)$. The formula (9), together with (10), gives $f(x, t') - f(x, t) > f(x', t') - f(x', t)$, so f violates increasing differences. $\square$

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

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

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

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

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

We show in the Appendix that Λ satisfies (F2), even though it is clear that it is *not* increasing in the strong set order (see Figure 3 on the right).

In the maxmin model, whenever $g(x, s)$ is increasing in s for all $x \in X$, one can assume, without loss of generality, that the belief correspondence Λ is upper comprehensive, i.e., if $\lambda \in \Lambda(t)$ and $\lambda' \succeq \lambda$ then $\lambda' \in \Lambda(t)$.[26] The following result (which we prove in the Appendix) states that when Λ is upper comprehensive, property (F2) remains a necessary and sufficient condition even if we only require the supermodularity of f for those functions g that are supermodular *and* increasing in s.

Proposition 6. *Suppose that correspondence $\Lambda : T \to \triangle_S$ has compact, convex, and upper comprehensive values. Then the following statements are equivalent.*

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

(ii) *The function $f(x, t)$, defined in (7), is supermodular in (x, t) for all supermodular functions g that are increasing in s.*[27]

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

[27] Remark 4.1 is also applicable to Proposition 6, i.e., statement (ii) is equivalent to $f(x, t)$ having *single crossing differences* in (x, t) for all supermodular functions g that are increasing in s. The proof of Remark 4.1 in the Appendix can be applied without modification to establish this claim.

Remark 4.4. Notice that the comparative statics problem is dramatically simplified if $g(x, s)$ is always increasing in s *and* $\Lambda(t)$ contains its infimum, i.e., a distribution $\underline{\lambda}(t)$ that is dominated by every other distribution in $\Lambda(t)$. Then $f(x, t) = \int_S g(x, s) d\underline{\lambda}(s)$ for all x and f is supermodular in (x, t) if $\underline{\lambda}(\cdot, t)$ increases with t. This is consistent with Proposition 6 because $\Lambda(t) = \{\lambda \in \triangle_S : \lambda \succeq \underline{\lambda}(t)\}$ satisfies (F2). However, there are natural examples of Λ obeying (F2) without $\Lambda(t)$ containing its infimum; see Example 7.

We conclude this subsection with three economic applications. Further applications are found in the Online supplement, where we apply our results to formulate conditions for monotone decision rules for ambiguity averse agents choosing in a dynamic context; this generalizes known results on monotone decision rules (see Hopenhayn and Prescott, 1992) for agents maximizing discounted expected utility.

Example 8 (Optimal savings)**.** Consider an agent who lives for two periods, with income m in period 1 and uncertain income s in period 2. The agent chooses saving $x \in [0, m]$ in period 1. Then

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

where u is the per-period utility, β is the discount rate, and r is the interest. The function g is increasing in s if u is increasing and it is submodular in (x, s) (equivalently, $g_{xs} \leq 0$) if u is concave. Suppose the agent has maxmin preferences of the form (7). Then the agent will be better off with higher t if Λ satisfies (F1), since g increases with s. Furthermore, since g is also submodular, f has increasing differences in (x, t) if Λ satisfies the stronger property (F2). This in turn guarantees that the agent saves less with higher t; formally, $\mathrm{argmax}_{x \in [0,m]} f(x, t)$ falls with t in the strong set order.

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

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

$$f(x,t) \ := \ \min\left\{\int_S u\big(xs + (m-x)r\big)d\lambda(s) : \lambda \in \Lambda(t)\right\}. \tag{12}$$

To capture the idea that a higher t represents greater optimism, we assume that correspondence Λ increases in t according to (F2). It follows that f has increasing differences in (x,t) if $g(x,s) := u\big(xs + (m-x)r\big)$ is supermodular. Assuming that u is strictly increasing, concave, and twice continuously differentiable, it is straightforward to check that g is supermodular if the coefficient of relative risk aversion of u is less than 1.[28] With this condition on u, f has increasing differences in (x,t) and (consequently) the investor's holding in the risky asset increases with t. This conclusion is valid even if the investor's preference has the α-maxmin form.[29]

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

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

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

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

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

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

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

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

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

It is straightforward to check that the function π is supermodular if κ is increasing, convex, and $\kappa(0) = 0$.[30] Proposition 5 guarantees that $f(x,t) = \min\left\{\int_S \pi(x,s)d\lambda(s) : \lambda \in \Lambda(t)\right\}$ is a supermodular function of (x,t) and therefore Π is supermodular in (x,t). Furthermore, Π has increasing differences in $\big((x,t),(-c,-a)\big)$. 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; formally, $\operatorname{argmax}_{(x,t)\in\mathbb{R}^2_+} \Pi(x,t,c,a)$ decreases with (c,a) in the strong set order (see Topkis, 1978).

4.2 Variational and multiplier preferences

Proposition 5 can be extended to cover a broader class of preferences. Maccheroni et al. (2006) introduce and axiomatize a generalization of the maxmin model, called *variational preferences*. In this model, the utility of some action x is

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

where c is a convex function. Loosely speaking, the agent's utility from action x is obtained by minimizing her expected utility over the set of all probability distributions; unlike the maxmin model where the agent is restricted to a subset of $\triangle_S$, any distribution in $\triangle_S$ could be 'picked' in the variational preferences model, though each distribution λ is associated with a different cost $c(\lambda)$.[31] Below, we parameterize the cost function c by $t \in T \subseteq \mathbb{R}$ and identify conditions under which the agent's utility has increasing differences in (x,t).

Proposition 7. *Let $c : \triangle_S \times T \to \mathbb{R}_+$ be a continuous and convex function on $\triangle_S$, for all $t \in T$. The following statements are equivalent.*

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

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

[31] For a discussion see Maccheroni et al. (2006) or Epstein and Schneider (2010).

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

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

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

$$f(x,t) \ := \ \min \left\{ \int_S g(x,s) d\lambda(s) + c(\lambda, t) : \lambda \in \triangle_S \right\}, \tag{13}$$

is supermodular for any supermodular function $g : X \times S \to \mathbb{R}$.[32]

Remark 4.5. To better understand condition (C), which may seem opaque initially, notice that it captures the change in the function c that leads to an upward revision in the agent's belief about the state. To be specific, suppose that λ_* and λ'_* are distributions that minimize $\int_S g(x,s)d\lambda(s)+c(\lambda,t)$ and $\int_S g(x,s)d\lambda(s)+c(\lambda,t')$ respectively, with $t' > t$. (C) guarantees that there are distributions μ_* and μ'_* such that $\lambda'_* \succeq \mu_*$, $\mu'_* \succeq \lambda_*$, and

$$\int_S g(x,s)d\lambda_*(s) + \int_S g(x,s)d\lambda'_*(s) + c(\lambda_*, t) + c(\lambda'_*, t') \geq \int_S g(x,s)d\mu_*(s) + \int_S g(x,s)d\mu'_*(s) + c(\mu_*, t) + c(\mu'_*, t').$$

Thus, μ_* also minimizes $\int_S g(x,s)d\lambda(s)+c(\lambda,t)$ and μ'_* minimizes $\int_S g(x,s)d\lambda(s)+c(\lambda,t')$. In other words, as t increases the distribution the agent uses to evaluate the utility of an action x shifts up in the sense of first order stochastic dominance (from λ_* to μ'_*).

The proof of Proposition 7 is in the Appendix. Note that condition (C) can be thought of as generalization of condition (F2) imposed on $\Lambda : T \to \triangle_S$. Indeed, given Λ, we define

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

Then c obeys (C) if and only if Λ obeys (F2) (while (13) reduces to the maxmin form (7) in this case). Below are two more examples of cost functions that satisfy property (C).[33]

Example 11 (Submodular cost and decreasing differences)**.** Let $c : \triangle_S \times T \to \mathbb{R}_+$ be a submodular function of λ that has decreasing differences in (λ, t). Then for for all λ, $\lambda' \in \triangle_S$ and t, $t' \in T$ with $t' > t$, we have

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

[32] Remark 4.1 is also applicable to Proposition 7, i.e., statement (ii) is equivalent to $f(x,t)$ having *single crossing differences* in (x,t) for all supermodular functions g that are increasing in s. The proof of Remark 4.1 in the Appendix can be applied to establish this claim, with f given by (13).

[33] Note that (C) restricts how $c(\lambda, t)$ varies jointly with λ and t; for a *fixed* t, it has no content.

and condition (C) holds, if we choose $\mu = (\lambda \wedge \lambda')$ and $\mu' = (\lambda \vee \lambda')$.

Example 12. Suppose that $\tilde{c} : \mathbb{R} \times T \to \mathbb{R}$ has decreasing differences in (m, t) and the cost function $c : \triangle_S \times T \to \mathbb{R}$ is evaluated by $c(\lambda, t) := \tilde{c}\left(\int_S h(s) d\lambda(s), t\right)$ for some increasing function $h : S \to \mathbb{R}$. In other words, the cost function depends only on the mean of the random variable h with respect to the distribution λ, and the parameter t. We claim that c satisfies (C). Let $t' > t$; take any λ, λ' in $\triangle_S$ and denote the mean of function h corresponding to each distribution by m, m', respectively. Suppose that $m' \geq m$; then there are distributions μ, μ' with means m, m', respectively, such that $\lambda' \succeq \mu$, $\mu' \succeq \lambda$, and $(1/2)\lambda + (1/2)\lambda' = (1/2)\mu + (1/2)\mu'$.[34] Since $c(\lambda, t) = c(\mu, t)$ and $c(\lambda', t') = c(\mu', t')$, we obtain (as required) $c(\lambda, t) + c(\lambda', t') = c(\mu, t) + c(\mu', t')$. If $m' < m$, then choose $\mu = \lambda'$ and $\mu' = \lambda$; since $\tilde{c}$ has decreasing differences in (m, t) we obtain

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

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

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

is the *relative entropy.*[35] Note that $d\lambda(s)$, $d\lambda^*(s, t)$ denote the probability of state s in the distribution λ, $\lambda^*(\cdot, t)$, respectively. This representation can be interpreted in the following manner. The decision maker has a belief over the states of the world given by a *reference* or *benchmark* distribution $\lambda^*(\cdot, t)$, but she is not completely confident that she is exactly correct. To accommodate this concern, the decision maker takes all distributions in $\triangle_S$ into account when evaluating her utility from a given action, though distributions further away from $\lambda^*(\cdot, t)$ cost more and are thus less likely to be the distribution that solves the minimization problem in (13).

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

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

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

Proposition 8. *The cost function* $c : \triangle_S \times T \to \mathbb{R}$*, given by* $c(\lambda, t) := \theta R\big(\lambda \| \lambda^*(\cdot, t)\big)$ *is submodular on* $\triangle_S$*, for all* $t \in T$ *and* $\theta > 0$*. Furthermore, if* $\lambda^*(\cdot, t)$ *is increasing in* t *with respect to the monotone likelihood ratio,*[36] *then* c *has decreasing differences in* (λ, t)*.*

We prove this result in the Appendix. This result tells us that so long as $\lambda^*(\cdot, t)$ is increasing in t with respect to the monotone likelihood ratio, then $c(\lambda, t)$ obeys condition (C) (see Example 11). By applying Proposition 7, we conclude that $f(x, t)$ is supermodular in (x, t) if $g(x, s)$ is supermodular in (x, s). This result captures the feature that as the agent revises her benchmark belief towards higher states, the cost function changes in a way that raises the marginal utility to her of taking higher actions.

In Examples 9 and 10, we gave economic applications of Proposition 5, which assume that the agent has maxmin utility. It is clear that, by appealing to Proposition 7, the conclusions in those examples will continue to hold, mutatis mutandi, if the agent has variational or, more specifically, multiplier preferences.

Appendix

Continuation of the proof to Theorem 1 We show that statement (iii) implies (i) by contradiction. Suppose Γ violates parallelogram property, so there is $t' \geq_T t$ and $x \in \Gamma(t), x' \in \Gamma(t')$ for which there is no $y \in \operatorname{co}\Gamma(t)$, $y' \in \operatorname{co}\Gamma(t')$ satisfying $x + x' = y + y'$ and $x'_K \geq y_K$, $y'_K \geq x_K$. Take any such x, x' and define

$$C \ := \ \Big\{(x - y', x' - y) \in \mathbb{R}^\ell \times \mathbb{R}^\ell : y \in \operatorname{co}\Gamma(t) \text{ and } y' \in \operatorname{co}\Gamma(t')\Big\}$$

and

$$D \ := \ \Big\{(d, d') \in \mathbb{R}^\ell \times \mathbb{R}^\ell : d + d' = 0 \text{ and } d'_K \geq 0\Big\}.$$

Clearly, both sets are closed, convex, and $C \cap D = \emptyset$. Moreover, since C is compact, one can show that $(D - C)$ is closed. By the strong separating hyperplane theorem, there are non-zero vectors $p, p' \in \mathbb{R}^\ell$ and a number b that satisfy

$$\sup\Big\{p \cdot c + p' \cdot c' : (c, c') \in C\Big\} \ < \ b \ < \ \inf\Big\{p \cdot d + p' \cdot d' : (d, d') \in D\Big\}. \qquad \text{(A1)}$$

[36] This requires that, for any $t' \geq t$, the ratio $d\lambda^*(s, t')/d\lambda^*(s, t)$ be increasing with s. This property implies $\lambda^*(\cdot, t') \succeq \lambda^*(\cdot, t)$, hence, it is stronger than the first order stochastic dominance.

Since $(0,0) \in D$, we have $b < 0$. Let $\epsilon_i \in \mathbb{R}^\ell_+$ be the vector with the i'th entry equal to 1 and zeros elsewhere, for $i = 1, 2, \ldots, \ell$. Given that $\alpha(-\epsilon_i, \epsilon_i) \in D$, for all $\alpha \geq 0$ and $i \in K$, we have $p'_K \geq p_K$. Since $\alpha(-\epsilon_i, \epsilon_i)$ belongs to D, for all numbers α and $i \notin K$, it must be that $p'_{-K} = p_{-K}$. The first inequality in (A1) gives $p \cdot x + p' \cdot x' < b + p \cdot y' + p' \cdot y$ for all $y' \in \Gamma(t')$ and $y \in \Gamma(t)$. Therefore,

$$\begin{aligned} f(p,t) + f(p',t') \; &\leq \; p \cdot x + p' \cdot x' \\ &< \; \min\left\{p \cdot z : z \in \Gamma(t')\right\} + \min\left\{p' \cdot z : z \in \Gamma(t)\right\} \; = \; f(p,t') + f(p',t), \end{aligned}$$

which contradicts our assumption that $f(p,t)$ has increasing differences in (p_K, t). □

Continuation of the proof of Theorem 2 We prove that (iii) $\Rightarrow$ (i). Toward contradiction, suppose that this property is violated for some $x \in \Gamma(t)$, $x' \in \Gamma(t')$. Define set C and D as in the the proof of Theorem 1. As previously, both sets are closed, convex, and $C \cap D = \emptyset$; furthermore, using the definition of D and the fact that C is bounded from above, one can show that $(D - C)$ is closed.[37] By the strong separating hyperplane theorem, there are non-zero vectors $\tilde{p}, \tilde{p}' \in \mathbb{R}^\ell$ and a number b such that

$$\sup\left\{\tilde{p} \cdot c + \tilde{p}' \cdot c' : (c,c') \in C\right\} \; < \; b \; < \; \inf\left\{\tilde{p} \cdot d + \tilde{p}' \cdot d' : (d,d') \in D\right\}. \tag{A2}$$

Since $(0,0) \in D$, we have $b < 0$. Given that the C is downward comprehensive, we have $\tilde{p}, \tilde{p}' \geq 0$. Finally, as in the proof of Theorem 1, we have $\tilde{p}'_K \geq \tilde{p}_K$ and $\tilde{p}_K = \tilde{p}'_K$.

We claim that there are vectors $p, p' \gg 0$ such that $p'_K \geq p_K$ and $p'_{-K} = p_{-K}$, that strictly separate the sets C and D.[38] Let $p := (\tilde{p} + \delta\mathbf{1})$ and $p' := (\tilde{p}' + \delta\mathbf{1})$, for some $\delta > 0$, where $\mathbf{1} \in \mathbb{R}^\ell$ is the unit vector. Clearly, we have $p', p \gg 0$ as well as $p'_K \geq p_K$ and $p'_{-K} = p_{-K}$. Given that $d + d' = 0$ for any $(d,d') \in D$, we obtain

$$p \cdot d + p' \cdot d' \; = \; \tilde{p} \cdot d + \tilde{p}' \cdot d' + \delta\mathbf{1} \cdot (d + d') \; = \; \tilde{p} \cdot d + \tilde{p}' \cdot d'$$

for any $\delta > 0$. Thus, $\inf\left\{p \cdot d + p' \cdot d' : (d,d') \in D\right\} = \inf\left\{\tilde{p} \cdot d + \tilde{p}' \cdot d' : (d,d') \in D\right\}$. Since C is bounded from above, we have $\sup\left\{p \cdot c + p' \cdot c' : (c,c') \in C\right\} < b$, for a sufficiently small

[37] Let d^n and c^n be sequences in D and C such that $(d^n - c^n)$ converges to a. If both d^n and c^n are convergent sequences, then the closedness of D and C will guarantee that $a \in D - C$ (as required). A problem arises when both d_n and c_n are divergent, but that is impossible. Indeed, with no loss of generality, suppose $c^n_1 \to -\infty$, which implies that $d^n_1 \to -\infty$, which (given the definition of D) implies that $d^n_{\ell+1} \to \infty$ which implies that $c^n_{\ell+1} \to \infty$, but that is impossible since C is bounded above.

[38] By $p \gg 0$ we mean that $p_i > 0$, for all $i = 1, 2, \ldots, \ell$.

$\delta > 0$. Thus (A2) holds, with $p \gg 0$ and $p' \gg 0$ taking the place of $\tilde{p}$ and $\tilde{p}'$. Re-tracing the proof that (iii) $\Rightarrow$ (i) in Theorem 1, we obtain $f(p,t) + f(p',t') < f(p,t') + f(p',t)$, contradicting our assumption that f has increasing differences in (p_K, t). $\square$

Proof of Remark 3.2 We show that H admits a selection that is normal in K. Without loss, suppose that $K = \{1, 2, \ldots, n\}$, for some $n \leq \ell$. Let $>_{lex}$ denote the lexicographic order[39] and $h(p,q) := \{x \in H(p,q) : x >_{lex} y, \text{ for all } y \in H(p,q)\}$. Since H is compact-valued, the function h is well-defined and $h(p,q) \in H(p,q)$, for all $(p,q) \in \mathbb{R}^{\ell}_{++} \times Q$. We show that $h_K(p,q') \geq h_K(p,q)$, for any $q' \geq q$. Toward contradiction, assume the opposite. By (ii), there is some $y \in H(p,q)$, $y' \in H(p,q')$ such that $h(p,q) - y = y' - h(p,q')$ and $y'_K \geq h_K(p,q)$, $h(p,q') \geq y_K$. Since $h_K(p,q) \neq y_K$, we have $h_K(p,q) >_{lex} y$. Thus, there is some $j \leq n$ such that $h_i(p,q) = y_i$, for all $i \leq j$, and $h_j(p,q) > y_j$. However, since $h_K(p,q) - y = y' - h_K(p,q')$, this implies $y'_i = h_i(p,q')$, for all $i \leq j$, and $y'_j > h_j(p,q')$. Hence, we have $y' >_{lex} h(p,q')$, which contradicts the definition of $h(p,q')$. $\square$

Proof of Proposition 2 We have already shown that (i) implies (ii). To prove the converse, note that, by Proposition 1, if F violates the parallelogram property for i then $C(p,q)$ violates increasing differences in (p_i, q); i.e., there is $p''_i > p'_i$ and $q'' > q'$ such that $C\big((p''_i, \hat{p}_{-i}), q''\big) - C\big((p''_i, \hat{p}_{-i}), q'\big) < C\big((p'_i, \hat{p}_{-i}), q''\big) - C\big((p'_i, \hat{p}_{-i}), q'\big)$, for some $\hat{p}_{-i}$. Choose R such that $R(q) = C(p''_i, \hat{p}_{-i}, q)$, for all $q < q''$, and $R(q) = C(p''_i, \hat{p}_{-i}, q'')$, for all $q \geq q''$. Given the assumptions on F, the cost C is strictly increasing in q; therefore, at price $(p''_i, \hat{p}_{-i})$ the firm is maximizing profit (which equals zero) at $q = q''$, but for all $q > q''$ the firm makes a loss. On the other hand, profit is *not* maximized at any $q \geq q''$ when $p = (p'_i, \hat{p}_{-i})$, which violates (ii).

To show that (i) implies (iii), let $\widehat{Q}(t) := \operatorname{argmax}_{q \in Q} \{R(q,t) - C(p,q)\}$. Since R has increasing differences, $\widehat{Q}(t)$ is increasing in t in the strong set order. Take any $\hat{x}' \in \widehat{X}(t')$. Then $F(\hat{x}') \in \widehat{Q}(t')$. For any $t'' > t'$, we have $q'' \geq F(\hat{x}')$, for some $q'' \in \widehat{Q}(t'')$. By Proposition 1, the factor demand for i is normal, and so there is some $\hat{x}'' \in H(p,q'')$, and thus $\hat{x}'' \in \widehat{X}(t'')$, such that $\hat{x}''_i \geq \hat{x}'_i$, as required. To show that (iii) implies (i), suppose that (i) fails; by Proposition 1, there is p, q' and q'' with $q' < q''$, and $x' \in H(p,q')$ such

[39] Recall that $x >_{lex} y$ if $x_i = y_i$, for all $i \leq j$, and $x_j > x_j$, for some $j \leq \ell$.

that $x_i' > x_i''$ for every $x'' \in H(p, q'')$.[40] We can choose $R(\cdot, t')$ and $R(\cdot, t'')$ satisfying increasing differences such that $q' = \widehat{Q}(t')$ and $q'' = \widehat{Q}(t'')$.[41] Then, we have $\hat{x}' \in \widehat{X}(t')$, but there is no $\hat{x}'' \in \widehat{X}(t'')$ such that $x_i' \leq \hat{x}_i''$, contradicting (iii). □

Proof of Proposition 4 To show that (i) ⇒ (ii), take any $t' \geq t$ and $\lambda' \in \Lambda(t')$. By (F1), there is some $\lambda \in \Lambda(t)$ such that $\lambda' \succeq \lambda$. Thus, for any increasing u,

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

Taking the minimum over the left term gives us the result.

To show (ii) ⇒ (i), suppose (F1) fails. Then there is $t' \geq t$ and $\lambda' \in \Lambda(t')$ such that $\lambda' \not\succeq \lambda$, for all $\lambda \in \Lambda(t)$. Let $V = \{y \in \mathbb{R}^\ell : y_i \geq \lambda'(s_i), \text{ for } i = 1, \ldots, \ell\}$. Since $V \cap \Lambda(t') = \emptyset$ and $(V - \Lambda(t'))$ is closed and convex, by the strong separating hyperplane theorem, $\min\left\{\sum_{i=1}^\ell \hat{p}_i y_i : y \in V\right\} > \max\left\{\sum_{i=1}^\ell \hat{p}_i \lambda(s_i) : \lambda \in \Lambda(t')\right\}$, for some $\hat{p} \in \mathbb{R}^\ell$. Given that V is upward comprehensive, $\hat{p} > 0$; furthermore, $\sum_{i=1}^\ell \hat{p}_i \lambda'(s_i) = \min\{\hat{p} \cdot y : y \in V\}$. Define $u : S \to \mathbb{R}$ by $u(s_1) = \hat{p}_1$ and $u(s_{i+1}) = u(s_i) + \hat{p}_{i+1}$, for $i = 1, \ldots, \ell$, which is an increasing function. Since $\int_S u(s) d\mu(s) = u(s_{\ell+1}) - \sum_{i=1}^\ell \hat{p}_i \mu(s_i)$, for any $\mu \in \triangle_S$,

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

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

Proof of Remark 4.1 Suppose that there is some supermodular function $\tilde{g}$ for which $\tilde{f}(\tilde{x}, t) := \min\{\int_S \tilde{g}(\tilde{x}, s) d\lambda(s) : \lambda \in \Lambda(t)\}$ violates increasing differences. In particular, there is some $x' \geq x$ and $t' \geq_T t$ such that

$$v \;:=\; f(x', t) - f(x, t) \;>\; f(x', t') - f(x, t').$$

[40] An analogous argument holds when $x'' \in H(p, q'')$ and $x_i' > x_i''$, for all $x' \in H(p, q')$.

[41] Let $R(\cdot, t')$ be a step function with a sufficiently large jump at q'; then the optimal output is uniquely at $q = q'$. (No higher output is optimal since C strictly increases with q.) Then choose $R(\cdot, t'')$ to be a step function that equals $R(\cdot, t')$ for all $q < q''$ and with another, sufficiently high, jump at $q = q''$.

Define the function g by $g(y,s) = \tilde{g}(y,s)$, for $y \leq x$, and $g(y,s) = \tilde{g}(y,s) - v$ otherwise. Clearly, g is supermodular, but f given by $f(\tilde{x},t) := \min\left\{\int_S g(\tilde{x},s)d\lambda(s) : \lambda \in \Lambda(t)\right\}$ violates single crossing differences since $0 = f(x',t) - f(x,t) > f(x',t') - f(x,t')$. □

Proof of Remark 4.2 Suppose $S = [a,b]$. Let $\{s_i^n\}_{i=0}^n$ be a sequence with $n+1$ terms such that $a = s_0^n < s_1^n < \ldots < s_{n-1}^n < s_n^n = b$. Since at each (x,t), function $g(x,\cdot)$ is the Riemann-Stieltjes integrable with respect to $\lambda \in \Lambda(t)$, we can choose $\{s_i^n\}_{i=0}^n$ so that

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

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

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

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

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

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

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

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

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

Proof of Proposition 6 Implication (i) $\Rightarrow$ (ii) follows from Proposition 5. We prove the converse in two steps. First, using Theorem 2 and an argument analogous to the one in the proof of Proposition 5, we show that the function f satisfies increasing differences only if the correspondence $\Gamma : T \to \mathbb{R}^\ell$, defined as

$$\Gamma(t) \;:=\; \Big\{y \in \mathbb{R}^\ell : y_i \geq -\lambda(s_i),\ \text{for all } i = 1,\ldots,\ell \text{ and some } \lambda \in \Lambda(t)\Big\}$$

satisfies the parallelogram property. This means that for any $t' \geq_T t$ and $\lambda \in \Lambda(t)$, $\lambda' \in \Lambda(t')$, there is $\mu \in \Lambda(t)$, $\mu' \in \Lambda(t')$, θ and $\theta' \in \mathbb{R}^\ell$ such that $\theta_i \leq \mu(s_i)$, $\theta'_i \leq \mu'(s_i)$, $\lambda(s_i) + \lambda'(s_i) = \theta_i + \theta'_i$, and $\theta_i \geq \lambda'(s_i)$ for all i. Therefore, Λ has the following property, which we shall refer to as ($\star$): for any $t' \geq_T t$ and $\lambda \in \Lambda(t)$, $\lambda' \in \Lambda(t')$, there is $\mu \in \Lambda(t)$, $\mu' \in \Lambda(t')$ such that $(1/2)\lambda + (1/2)\lambda' \succeq (1/2)\mu + (1/2)\mu'$ and $\lambda' \succeq \mu$.

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

Proof of Proposition 7 Let $t'' >_T t'$ and let M satisfy $M > \max\left\{c(\lambda, t'), c(\lambda, t'')\right\}$, for all $\lambda \in \triangle_S$. The correspondence $\Gamma : T' \to \mathbb{R}^{\ell+1}$, where $T' = \{t', t''\}$, is defined by

$$\Gamma(t) \;:=\; \Big\{y \in \mathbb{R}^{\ell+1} : y_i = -\lambda(s_i), \text{ for } i = 1, 2, \ldots, \ell, \text{ and } y_{\ell+1} \in \left[c(\lambda, t), M\right] \text{ for } \lambda \in \triangle_S\Big\}.$$

Since c is convex, Γ is convex-valued; using this fact, it is straightforward to check that Γ satisfies the parallelogram property for $K = \{1, 2, \ldots, \ell\}$ if and only if the function c obeys (C). Suppose that c satisfies (C); given a supermodular function g, we define

$p', p \in \mathbb{R}^\ell$ by $p'_i := g(x', s_{i+1}) - g(x', s_i)$ and $p_i := g(x, s_{i+1}) - g(x, s_i)$, for $i = 1, \ldots, \ell$. The supermodularity of g guarantees that $p' \geq p$ if $x' \geq x$. By Theorem 1 and the integration formula (8), for any $x' \geq x$ and $t' \geq_T t$, we obtain

$$f(x,t') - f(x,t) \; = \; \min\Big\{(p,1)\cdot y : y \in \Gamma(t')\Big\} - \min\Big\{(p,1)\cdot y : y \in \Gamma(t)\Big\}$$
$$\leq \; \min\Big\{(p',1)\cdot y : y \in \Gamma(t')\Big\} - \min\Big\{(p',1)\cdot y : y \in \Gamma(t)\Big\} \; = \; f(x',t') - f(x',t).$$

We conclude that (i) $\Rightarrow$ (ii).

We prove the converse by contradiction. Suppose c violates (C) and so Γ violates the parallelogram property for $K = \{1, 2, \ldots, \ell\}$. By Theorem 1, function $\tilde{f} : \mathbb{R}^{\ell+1} \times T' \to \mathbb{R}$, $\tilde{f}\big((\tilde{p}, q), t\big) := \min\big\{(\tilde{p}, q)\cdot y : y \in \Gamma(t)\big\}$, must violate increasing differences in $(\tilde{p}, t)$, i.e., there is $p, p' \in \mathbb{R}^\ell$, $t, t' \in T$, and $q \in \mathbb{R}$ such that $p' \geq p$, $t' >_T t$ and

$$\tilde{f}\big((p',q),t\big) - \tilde{f}\big((p,q),t\big) \; > \; \tilde{f}\big((p',q),t'\big) - \tilde{f}\big((p',q),t'\big).$$

If $q \leq 0$, then $\tilde{f}\big((p,q),t\big) = \tilde{f}\big((p,q),t'\big)$ and $\tilde{f}((p',q),t) = \tilde{f}((p',q),t')$, so we only need to consider $q > 0$. Given this, we can assume with no loss of generality that $q = 1$, so that

$$\tilde{f}\big((\tilde{p},1),t\big) \; = \; \min\left\{\sum_{i=1}^{\ell} \tilde{p}_i\big[-\lambda(s_i)\big] + c(\lambda,t) : \lambda \in \Delta_S\right\}.$$

Define the function $g : X \times S \to \mathbb{R}$ as in (11). Using formula (8) again, we obtain

$$f(x',t) - f(x,t) \; = \; \tilde{f}\big((p',1),t\big) - \tilde{f}\big((p,1),t\big) \; > \; \tilde{f}\big((p',1),t'\big) - \tilde{f}\big((p,1),t'\big) \; = \; f(x',t') - f(x,t')$$

which means that f violates increasing differences. $\square$

Proof of Proposition 8 It suffices to show that $R\big(\lambda \| \lambda^*(\cdot,t)\big)$ is submodular in λ (for each t) and has decreasing differences in (λ, t). To prove the former, let λ, $\lambda' \in \triangle_S$ and denote $\mu' = (\lambda \vee \lambda')$ and $\mu = (\lambda \wedge \lambda')$. To abbreviate the notation, let p_i, p'_i, q_i, q'_i be the probability of state s_i, for all $i = 1, 2, \ldots, (\ell+1)$, corresponding to the cumulative distribution λ, λ', μ, μ', respectively. $R\big(\lambda \| \lambda^*(\cdot,t)\big)$ is submodular in λ if, for all i,

$$p_i \ln p_i + p'_i \ln p'_i - \big[p_i + p'_i\big] \ln d\lambda^*(s_i,t) \; \geq \; q_i \ln q_i + q'_i \ln q'_i - \big[q_i + q'_i\big] \ln d\lambda^*(s_i,t). \quad \text{(A3)}$$

Clearly, this inequality holds for $i = 1$. Consider $i > 1$. With no loss of generality, let $\mu(s_{i-1}) = \lambda(s_{i-1})$ and $\mu'(s_{i-1}) = \lambda'(s_{i-1})$. Consider two cases. Assume that (i) $p_i' + \lambda'(s_{i-1}) \leq p_i + \lambda(s_{i-1})$, so that $\mu(s_i) = \lambda(s_i)$ and $\mu'(s_i) = \lambda'(s_i)$. Then $q_i = p_i$ and $q_i' = p_i'$ and (A3) is satisfied with equality. Suppose, instead, that (ii) $p_i' + \lambda'(s_{i-1}) > p_i + \lambda(s_{i-1})$, which implies $\mu(s_i) = \lambda'(s_i)$ and $\mu'(s_i) = \lambda(s_i)$. Let $\delta = \lambda(s_{i-1}) - \lambda'(s_{i-1})$ and notice that $0 \leq \delta < p_i' - d\lambda(s_i)$. Since $q_i = p_i' - \delta$ and $q_i' = p_i + \delta$,

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

where the last inequality follows from the convexity of the map $z \to z \log z$. Thus, condition (A3) holds for all i and thus $R\big(\lambda \| \lambda^*(\cdot, t)\big)$ is submodular in λ.

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

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

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

References

Afriat, S. N. (1967): "The construction of a utility function from expenditure data," *International Economic Review*, 8, 67–77.

Alarie, Y., C. Bronsard, and P. Ouellette (1990): "Preferences and normal goods : A necessary and sufficient condition," *Journal of Economic Theory*, 51, 423–430.

Bergstrom, T., L. Blume, and H. Varian (1986): "On the private provision of public goods," *Journal of Public Economics*, 29, 25–49.

Bilancini, E. and L. Boncinelli (2010): "Preferences and normal goods: An easy-to-check necessary and sufficient condition," *Economics Letters*, 108, 13–15.

Blundell, R., M. Browning, and I. Crawford (2003): "Nonparametric Engel curves and revealed preference," *Econometrica*, 71, 205–240.

Blundell, R., P.-A. Chiappori, and C. Meghir (2005): "Collective labor supply with children," *Journal of Political Economy*, 113, 1277–1306.

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

Cherchye, L., T. Demuynck, and B. De Rock (2018): “Normality of demand in a two-goods setting,” *Journal of Economic Theory*, 173, 361–382.

Chipman, J. (1977): “An empirical implication of Auspitz-Lieben-Edgeworth-Pareto complementarity,” *Journal of Economic Theory*, 14, 228–231.

Dziewulski, P. and J. K. Quah (2021): “Supplement to: 'Comparative statics with linear objectives: normal demand, monotone marginal costs, and ranking multi-prior beliefs',” Online supplement.

Epstein, L. G. and M. Schneider (2010): “Ambiguity and asset markets,” *Annual Review of Financial Economics*, 2, 315–346.

Fisher, F. M. (1990): “Normal goods and the expenditure function,” *Journal of Economic Theory*, 51, 431–433.

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

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

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

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

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

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic theory*, Oxford University Press.

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

Nachbar, J. H. (2002): “General equilibrium comparative statics,” *Econometrica*, 70, 2065–2074.

Quah, J. K.-H. (2003): “Market demand and comparative statics when goods are normal,” *Journal of Mathematical Economics*, 39, 317–333.

——— (2007): “The comparative statics of constrained optimization problems,” *Econometrica*, 75, 401–431.

Rockafellar, T. R. (1970): *Convex analysis*, Princeton University Press.

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

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

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

Topkis, D. M. (1978): "Minimizing a submodular function on a lattice," *Operations Research*, 26, 305–321.

Ui, T. (2015): "Ambiguity and risk in global games," Working paper.

Supplement to: "Comparative statics with linear objectives: normal demand, monotone marginal costs, and ranking multi-prior beliefs"

Paweł Dziewulski* John K.-H. Quah†

February, 2021

Abstract

This supplement contains additional results related to Dziewulski and Quah (2021). These notes should be read in conjunction with the main paper.

We devote the first section of this supplement to an application of Proposition 5 in the main paper to dynamic programming under ambiguity. In the subsequent section we discuss proofs of the claims made in Examples 1–3 in the main paper. Throughout we employ the notation introduced in the main paper.

B.1 Dynamic programming under ambiguity

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

Consider an agent who faces a stochastic control problem where X and S are the sets of endogenous and exogenous state variables, respectively. To keep the exposition

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

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

simple, we shall assume that X is a sublattice of a Euclidean space and S is a subset of another Euclidean space. The evolution of s over time follows a Markov process with the transition function λ. The agent's problem can be formulated in the following way (see Stokey et al., 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, which we denote by $B(x_\tau, s_\tau) \subseteq X$. The single-period return is given by the function $F : X \times S \times X \to \mathbb{R}$; $F(x, s, y)$ is the payoff when (x, s) is the state variable in period τ and y is the endogenous state variable in period $\tau + 1$ chosen in period τ. Finally, we assume that 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 F and the continuity of B — this problem admits a recursive representation, where $v = v^*$ is the unique solution to the Bellman equation

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

where $\lambda(\cdot, s)$ is a cumulative probability distribution over states of the world in the following period, conditional on the current state s.[1] Furthermore, the set

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

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

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

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

[1] See Theorem 9.6 in Stokey et al. (1989) for details.

is supermodular in y and has increasing differences in $\big(y,(x,s)\big)$. By Theorem 6.1 in Topkis (1978), $\Phi(x,s)$ is a compact sublattice of X and is increasing in (x,s).[2] This in turn guarantees the existence of the greatest optimal selection

$$\phi(x,s) \;:=\; \Big\{y \in \Phi(x,s) : y \geq_X z, \text{ for all } z \in \Phi(x,s)\Big\},$$ [3]

that is increasing and Borel measurable. Lastly, the policy function ϕ induces a Markov process on $X \times S$, where, for measurable sets $Y \subseteq X$ and $T \subseteq S$, the probability of $Y \times T$ conditional on (x,s) is the probability of T conditional on s if $\phi(x,s) \in Y$, and it is zero otherwise. HP make use of the monotonicity of ϕ to guarantee that this Markov process has a stationary distribution.[4]

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

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

[2] Condition (ii) on B guarantees that $B(x,s)$ is sublattice of X and that it increases with (x,s) in the strong set order. Given with the properties on f, we know that $\Phi(x,s)$ is a sublattice and that it increases with (x,s); this follows from Theorem 6.1 in Topkis (1978).

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

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

possible value of the exogenous variable in the following period, after observing s in the current period, is given by a *set* of distribution functions $\Lambda(s)$; this set depends on the current realization of the exogenous variable and is time-invariant. The set $\mathcal{M}$, given an initial value s_0, is obtained by concatenating the transition probabilities. Therefore, the probability of a path $(s_1, s_2, s_3, \ldots)$ is $\prod_{i=1}^{\infty} p_i$, where p_1 is the probability of s_1 for some distribution in $\Lambda(s_0)$, p_2 is the probability of s_2 for some distribution in $\Lambda(s_2)$, etc.

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

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

where $(Av)(y,s) = \min\left\{ \int_S v(y,s)d\lambda(s) : \lambda \in \Lambda(s)\right\}$ (see Iyengar, 2005). Furthermore, the problem is *well-behaved* in the sense defined at the beginning of this section.

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 5 tells us that this holds if the beliefs correspondence Λ satisfies (F2).

Proposition B.1. *Consider a well-behaved optimal control problem where X, $S \subseteq \mathbb{R}$, with X compact and S finite. Let $F(x,s,y)$ be supermodular in (x,s,y), $\Lambda : S \to \triangle_S$ satisfy (F2), and the graph of $B : X \times S \to X$ be a sublattice; then the value function $v^*(x,s)$ is supermodular. Furthermore, the correspondence $\Phi : X \times S \to \mathbb{R}$, where*

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

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

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

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

is a continuous function on $X \times S$ and $\mathscr{T}^n v$ converges uniformly to v^* as $n \to \infty$. By Proposition 5 in the main paper, whenever function v is supermodular, then so is Av.

This implies that $F(x,s,y) + \beta(Av)(y,s)$ is supermodular over $X \times S \times X$. Given that the graph of correspondence B is a sublattice, by Theorem 4.3 in Topkis (1978), the function $\mathscr{T}v$ is supermodular in (x,s). Since supermodularity is preserved under uniform convergence, we conclude that $v^* = \mathscr{T}v^*$ is a supermodular function of (x,s). The set $\Phi(x,s)$ consists of elements y that maximize $F(x,s,y) + \beta(Av^*)(x,s)$ over $B(x,s)$. Since the objective function is supermodular, while values of correspondence B are complete sub-lattices of X, by Theorem 6.1 in Topkis (1978), set $\Phi(x,s)$ is a complete sub-lattice of X. Furthermore, since B increases over $X \times S$ in the strong set order, so does Φ. As the problem is well-behaved, $\Phi(x,s)$ admits the greatest selection $\phi(x,s)$ and this selection is increasing. That ϕ is Borel measurable follows from standard arguments (see HP). □

Below we discuss an application of this result.

Example B.1. Consider the following dynamic optimization problem of a firm. In each period, the firm collects revenue $\pi(x,s)$, where $s \in S$ denotes the realized exogenous state of the world and $x \in \mathbb{R}_+$ is the level of capital stock currently available to the firm. Once s is revealed to the firm and the revenue collected, the firm may invest $a \in [0,K]$ at a cost $c(a)$, K being a finite positive number. With this investment, capital stock in the next period is $y = \delta x + a$, where $\delta \in [0,1]$ denotes the fraction of non-depreciated capital. Therefore, the dividend in each period is

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

where the firm chooses y from the interval $B(x,s) = [\delta x, \delta x + K]$. We know from HP that if the firm is an expected utility maximizer and the optimal control problem is well-behaved, the firm has a policy function that is increasing in (x,s) under these additional conditions: the transition *function* $\Lambda : S \to \triangle_S$ is increasing with respect to first order stochastic dominance and F is supermodular; the latter is guaranteed if π is supermodular and c is concave. Proposition B.1 goes further by saying that this remains true if the firm has a maxmin preference, so long as the transition *correspondence* Λ satisfies (F2).

B.2 Technologies generating normal factor demand

Here we present proofs of the claims made in Examples 1–3 in the main paper.

B.2.1 Example 1: Homothetic production

We show a direct proof that any homothetic function satisfies parallelgoram property. Recall that the function $F : X \to \mathbb{R}$ define over a cone $X \subseteq \mathbb{R}^\ell_+$ is *homothetic* if for any non-zero $x, x' \in X$ there is a $\lambda > 0$ such that $F(x') = F(\lambda x)$ and $F\big((1/\lambda)x'\big) = F(x)$.

To show that any such technology satisfies parallelogram property, take any $q' \geq q$ and $x, x' \in X$ such that $F(x) \geq q$, $F(x') \geq q'$. By construction, there is $z \leq x$, $z' \leq x'$ such that $F(z) \geq q$, $F(z') \geq q'$. Whenever $x \geq x'$ or $F(x) \geq F(x')$, set $y = x'$ and $y' = x$. If $x' > x$, let $y = x$ and $y' = x'$. In either case the required condition holds.

If $F(x') > F(x)$ and x, x' are unordered, they must both be non-zero. Since F is homothetic, there is some $\lambda > 0$ such that $F(\lambda x) = F(x')$ and $F\big((1/\lambda)x'\big) = F(x)$. If $\lambda \leq 1$, set $y = x'$ and $y' = x$. Since $F(x') \geq q' \geq q$, and $x > \lambda x$, we have $y \in U(q)$ and $y' \in U(q')$. For $\lambda > 1$, let $y = (1/\lambda)z'$ and $y' = x + x' - y$. Obviously, we have $y \in U(q)$. Furthermore, since $y' = x + x' - (1/\lambda)x' = (1/\lambda)(\lambda x) + (1 - 1/\lambda)x'$, where $x', (\lambda x) \in U(q')$, we have $y' \in \operatorname{co} U(q')$. This suffices for F to satisfy the parallelogram property. $\square$

B.2.2 Example 2: $\mathcal{C}$-flexible order

We divide this subsection into two parts.

Sufficient conditions for $\mathcal{C}$-flexible order Let $X \subseteq \mathbb{R}^\ell$ to be a convex lattice. In the main paper, we claimed that a function $F : X \to \mathbb{R}$ is increasing in the $\mathcal{C}$-flexible order for $K \subseteq \{1, \ldots, \ell\}$ if F is continuous, increasing, supermodular, and concave in x_{-i}, for all $i \in K$. This result can be found in Quah (2007); we provide a proof here for easy reference.

Take any $q' \geq q$ and $x, x' \in X$ such that $x'_K \not\geq x_K$ and $F(x) \geq q$, $F(x') \geq q'$. We show that there is a $\lambda \in [0, 1]$ satisfying $F\big(\lambda x' + (1-\lambda)(x \wedge x')\big) \geq q$, $F\big(\lambda x + (1-\lambda)(x \vee x')\big) \geq q$. This suffices for F to be increasing in the $\mathcal{C}$-flexible order for K.

Consider two cases. (i) If $F(x \wedge x') \geq q$, set $\lambda = 0$. By monotonicity of F, we have $F(x \wedge x') \geq F(x') \geq q'$. Next, let (ii) $F(x \wedge x') < q$. Since $q \leq q' \leq F(x')$, by continuity of F there is some $\lambda \in [0, 1]$ such that $F\big(\lambda x' + (1 - \lambda)(x \wedge x')\big) = q$. Denote $v = x' - (x \wedge x') = (x \vee x') - x$, which is a positive vector. Since $x'_K \not\geq x_K$, there is some

$i \in K$ such that $v_i = 0$. In particular, we obtain

$$q' - q \;\leq\; F(x') - F\big(\lambda x' + (1-\lambda)(x \wedge x')\big) \;=\; F(x') - F\big(x \wedge x' + \lambda v\big) \\ \leq\; F(x \vee x') - F\big(x + \lambda v\big) \;\leq\; F\big((x \vee x') - \lambda v\big) - F(x),$$

where the second inequality follows from supermodularity of F and the third is implied by the fact that F is concave in x_{-i} and $v_i = 0$.[5] Therefore, since $F(x) \geq q$, it must be that $q' \leq F\big((x \vee x') - \lambda v\big) = F\big(\lambda x + (1-\lambda)(x \vee x')\big)$. □

Decreasing marginal rates of substitution We revisit the second part of Example 2 in the main paper. Suppose that $X := X_1 \times X_2$, where X_1, X_2 are intervals of $\mathbb{R}$, and function $F : X \to \mathbb{R}$ is continuously differentiable and strictly increasing.

First, we show that if, for any $x_i \in X_i$, the marginal rate of substitution

$$-\left[\frac{\partial F}{\partial x_i}(x_1, x_2) \Big/ \frac{\partial F}{\partial x_j}(x_1, x_2)\right]$$

is a decreasing function of x_j, where $j \neq i$, then F increases in the $\mathcal{C}$-flexible order for $K = \{i\}$. We prove this result for $i = 1$, without loss of generality. Take any x, $x' \in X$ such that $x_1 > x'_1$. If $F(x) \geq F(x')$, then for any $q' \geq q$ such that $F(x) \geq q$ and $F(x') \geq q'$, the required condition holds for $\lambda = 1$. Suppose that $F(x) < F(x')$. By strict monotonicity of F, this holds only if $x_2 < x'_2$. We need to show that, for any $q' \geq q$ such that $F(x) \geq q$ and $F(x') \geq q'$, there is some $\lambda \in [0, 1]$ satisfying

$$F\big(x'_1, \lambda x'_2 + (1-\lambda)x_2\big) \geq q \quad \text{and} \quad F\big(x_1, \lambda x_2 + (1-\lambda)x'_2\big) \geq q'.$$

Define function $g_x : [x'_1, x_1] \to \mathbb{R}$ by $g_x(z) := \big\{z' \in \mathbb{R} : F(z, z') = F(x)\big\}$, which is well-defined and continuously differentiable. Let $g_{x'} : [x'_1, x_1] \to \mathbb{R}$ be defined analogously. Given monotonicity of the function F, we have $g_{x'}(z) \geq g_x(z)$, for all $z \in [x'_1, x_1]$. By the implicit function theorem and the condition on the MRS, we obtain

$$\frac{dg_x}{dz}(z) \;=\; -\frac{\frac{\partial F}{\partial x_1}\big(z, g_x(z)\big)}{\frac{\partial F}{\partial x_2}\big(z, g_x(z)\big)} \;\geq\; -\frac{\frac{\partial F}{\partial x_1}\big(z, g_{x'}(z)\big)}{\frac{\partial F}{\partial x_2}\big(z, g_{x'}(z)\big)} \;=\; \frac{dg_{x'}}{dz}(z).$$

Set $\lambda \in [0, 1]$ such that $F\big(x'_1, \lambda x'_2 + (1-\lambda)x_2\big) = F(x)$. Since $F(x') > F(x)$ and the function F is strictly increasing and continuous, such a number exists. Moreover, we have

[5] We are making use of the fact that when f is concave in direction v, we have $f(x - v) - f(x) \geq f(x - v - tv) - f(x - tv)$, for any $x \in X$ and scalar $t > 0$.

$g_x(x_1') = \big[\lambda x_2' + (1-\lambda)x_2\big]$. Given that $g_x(x_1) = x_2$ and $g_{x'}(x_1') = x_2'$, we obtain

$$x_2 - \big[\lambda x_2' + (1-\lambda)x_2\big] \;=\; g_x(x_1) - g_x(x_1') \;=\; \int_{x_1'}^{x_1} \tfrac{dg_x}{dz}(z)dz$$
$$\geq \int_{x_1'}^{x_1} \tfrac{dg_{x'}}{dz}(z)dz \;=\; g_{x'}(x_1) - g_{x'}(x_1') \;=\; g_{x'}(x_1) - x_2',$$

which implies $g_{x'}(x_1) \leq \big[\lambda x_2 + (1-\lambda)x_2'\big]$, and so $F\big(x_1, \lambda x_2 + (1-\lambda)x_2'\big) \geq F(x')$. This suffices to show that the function F increases in the $\mathcal{C}$-flexible set order for $i = 1$.

One can show that, if F is (in addition) quasiconcave, then the restriction imposed on the marginal rate of substitution is also necessary for the function to increases in the $\mathcal{C}$-flexible set order. In fact, since any such function satisfies parallelogram property, it is enough to show that decreasing MRS is necessary for the function to satisfy the latter.

We show it by contradiction. Suppose that the function F satisfies parallelogram property for $K = \{1\}$, but the monotonicity of MRS is violated. By monotonicity and differentiability of F, one can always find some $z, z' \in X$ such that

$$p \;:=\; \frac{\frac{\partial F}{\partial x_1}\big(z_1, z_2\big)}{\frac{\partial F}{\partial x_2}\big(z_1, z_2\big)} \;>\; \frac{\frac{\partial F}{\partial x_1}\big(z_1', z_2'\big)}{\frac{\partial F}{\partial x_2}\big(z_1', z_2'\big)} \;=:\; p',$$

where $z_1 > z_1'$ and $F(z') \geq F(z)$. Moreover, we have $p, p' > 0$. Let $q = F(z)$ and $q' = F(z')$. Since F is quasiconcave, we have $(p, 1) \cdot (y - z) \geq 0$, for all y such that $F(y) \geq q$, and $(p', 1) \cdot (y' - z') \geq 0$ if $F(y') \geq q'$. To show that the parallelogram property is violated, note that, for any y, y' such that $F(y) \geq q$, $F(y') \geq q'$, and $y + y' = z + z'$,

$$p'(y_1' - z_1') \;\geq\; z_2' - y_2' \;=\; y_2 - z_2 \;\geq\; p(z_1 - y_1).$$

However, since $p > p'$ and $z_1 - y_1 = y_1' - z_1'$, this can hold only if $y_1 = z_1$ and $z_1' = y_1'$. Since $z_1 > z_1'$, for any such vectors y, y' it must be that $y_1 > z_1'$. Given that the function F is quasiconcave, this suffices for the parallelogram property to fail. □

B.2.3 Example 3: Aggregating paralellogram property

In this section we prove the claim in Example 3. Let $f^j : \mathbb{R}^\ell_+ \to \mathbb{R}$, for $j = 1, 2 \ldots, n$, be a collection of continuous and concave functions that obey the parallelogram property for $K \subseteq \{1, \ldots, \ell\}$. In particular, continuity guarantees that the range Q^j of f^j is an interval, for all $j = 1, \ldots, n$. Moreover, let $G : \times_{j=1}^n Q^j \to \mathbb{R}$ be an aggregator that is increasing in the $\mathcal{C}$-flexible order. We show directly that the function

$$F(x) \;:=\; \max\left\{G\Big(f^1(y^1), f^2(y^2), \ldots, f^n(y^n)\Big) : x \geq \sum_{j=1}^{n} y^j\right\}$$

generates a factor demand that is normal in K. By Proposition 1 in the main paper, this is equivalent to F satisfying the parallelogram property for K and also equivalent to its cost function having increasing differences in (p_K, q).

Let C, c^j denote cost functions corresponding to production functions F, f^j, for all $j = 1, \ldots, n$. Given that f^j is concave, each function c^j is convex in q. Note that

$$C(p,q) \;=\; \min\left\{\sum_{j=1}^{n} c^j(p,q^j) : G(q^1,\ldots,q^n) \geq q\right\}. \tag{B.1}$$

Suppose $x' \in H(p,q')$ and let $q'' > q'$. To show directly that H is normal in K, we must show that there is $x'' \in H(p,q'')$ such that $x''_K \geq x'_K$. (The proof that there is $x'' \in H(p,q'')$ such that $x''_K \leq x'_K$ if $q'' < q'$ is similar.)

Since $x' \in H(p,q')$, there is $\tilde{y}^j$ for all j, such that $\sum_{j=1}^{N} \tilde{y}^j = x'$ and

$$(\tilde{q}^j)_{j=1}^{N} \in \operatorname{argmin}\left\{\sum_{j=1}^{n} c^j(p,q^j) : G(q^1,\ldots,q^n) \geq q\right\},$$

where $\tilde{q}^j = f^j(\tilde{y}^j)$. Since G is increasing in the $\mathcal{C}$-flexible set order and the map from $q = (q^1, q^2, \ldots, q^N)$ to $\sum_{j=1}^{n} c^j(p,q^j)$ is an additive and convex function of q, we know that there exists $(\hat{q}^j)_{j=1}^{N}$ such that $\hat{q}^j \geq \tilde{q}^j$ for all j and

$$(\hat{q}^j)_{j=1}^{N} \in \operatorname{argmin}\left\{\sum_{j=1}^{n} c^j(p,q^j) : G(q^1,\ldots,q^n) \geq q''\right\}$$

(see Quah, 2007). Now since each f^j satisfies parallelogram property for K and is concave, it generates a factor demand that is normal in K (see Proposition 1 and Remark 3.2 in the main paper). Therefore, for each j, there is $\hat{y}^j$, the factor demand generated by f^j at price p and output $\hat{q}^j$, such that $\hat{y}^j_K \geq \tilde{y}^j_K$. Then $x'' = \sum_{j=1}^{N} \hat{y}^j \in H(p,q'')$ and $x''_K \geq x'_K$.
□

As an application of this aggregation result, consider the following examples, where $G(q^1,q^2) = q^1 + q^2$:

(i) $F(x_1,x_2) := \max\left\{\sqrt{y_1\, y_2} + \sqrt{z_1 + z_2} : x_i \geq y_i + z_i,\ \text{ for } i = 1,2\right\}$, for $X = \mathbb{R}^2_+$;

(ii) $F(x_1,x_2,x_3,x_4) := \sqrt{x_1\, x_2} + \sqrt{x_3 + x_4}$, for $X = \mathbb{R}^4_+$;

(iii) $F(x_1, x_2, x_3) := \max \left\{ \sqrt{x_1 \, y} + \sqrt{x_3 + z} : x_2 \geq y + z \right\}$, for $X = \mathbb{R}^3_+$.

In all three cases, the plant level production functions, i.e., the square root of the product of inputs and the square root of the sum of inputs, are homothetic and concave functions. In case (i), both factors of production are used in both plants; in case (ii), the two plants produce with completely distinct factors; in case (iii), one factor (factor 2) is used by both plants. In all of these cases, overall factor demand is normal and it follows from Proposition 1 in the main paper that F satisfies the parallelogram property.

It is worth emphasizing that in this result, it is the parallelogram property that is preserved by our aggregation formula and not necessarily some other stronger properties that guarantee normality. For example, notice that the (aggregate) production functions F in (i), (ii), and (iii) are not in general homothetic, even though the plant level production functions do satisfy that property. So these examples tell us that the homothetic property, even when it holds for f^j for all j, need not be preserved by our aggregation formula. We also know that the $\mathcal{C}$-flexible set order is not preserved by this aggregation. For example, consider the case of $F(x_1, x_2, x_3) = \min\{x_1, x_2 + x_3\}$. Note that the aggregator $G(q^1, q^2) = \min\{q^1, q^2\}$ and also the functions $f^1(x_1) = x_1$ and $f^2(x_2, x_3) = x_2 + x_3$ are all increasing in the $\mathcal{C}$-flexible set order, and so an application of our result tells us that F satisfies the parallelogram property. (We can also see this directly, of course, since F is homogeneous of degree 1.) However, the function F is *not* increasing in the $\mathcal{C}$-flexible set order (see Example 2 and, in particular, footnote 19, in the main paper).

References

Dziewulski, P. and J. K. Quah (2021): "Comparative statics with linear objectives: On normal demand, monotone marginal costs, and ranking multi-prior beliefs," Working paper.

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

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

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

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

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

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

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

Topkis, D. M. (1978): "Minimizing a submodular function on a lattice," *Operations Research*, 26, 305–321.