Stochastic Games and Monotone Comparative Statics

Gregorio Curello

St Cross College

Doctorate of Philosophy in Economics

Extended Abstract

In Chapter 1, I study decision problems under uncertainty involving the choice of a rule mapping states into actions. I show that for any rule, there exists an increasing rule inducing larger expected payoffs for all payoff functions that are supermodular in action and state. I present five applications. The main ones are

1. A planner implementing a subsidy based on household data, subject to a budget constraint, may improve any schedule failing to transfer more to households with data indicating larger returns, even without knowing their specific utility functions.

2. A monopolist price-discriminating in a market where wealthier buyers are less price-sensitive should charge them more, even if there is a positive demand externality. Shifting high prices onto wealthier buyers while preserving the price distribution is profitable.

3. If an insurer selling a single product to several buyers has access to a signal of the shocks affecting them, she should offer an insurance that pays more when signals indicating larger shocks are observed, even if the shocks and the buyers' degrees of risk-aversion are heterogeneous.

In Chapters 2 and 3, I analyse a new class of dynamic public-good games arising among innovators of a common technology who share improvements with one other. Innovators could be firms part of an R&D alliance or an informal disclosure agreement. The game is novel as firms induce randomly-sized increments in the stock (innovations), and their payoffs are a general function of private effort (R&D) and the current stock (the technology level).

In Chapter 2, I characterise the social-welfare benchmark, and compute the unique symmetric Markov equilibrium of the game. I show that, in the welfare benchmark, agents decrease their effort and are better-off as the stock grows. In contrast, innovations may be detrimental in the symmetric equilibrium. Namely, if the opportunity cost of effort does not depend on the current stock, then any innovation is beneficial. However,

if the cost of effort is sufficiently increasing in the stock, and there are sufficiently many agents, moderate innovations are detrimental early on.

In Chapter 3, I characterise the payoff-maximising equilibrium of the two-player game with binary effort, when strategies may depend on time and the past trajectory of the stock. I show that the equilibrium is asymmetric and non-Markov. In particular, only one agent exerts effort after the stock exceeds a certain cutoff and, before this occurs, agents face an equal chance of being the ones to do so. Moreover, I extend the game by allowing agents to conceal the innovations they obtain for an arbitrary period of time, foregoing their benefits. I derive necessary and sufficient conditions for the symmetric equilibrium of the baseline model with full disclosure to carry over to the game with concealment. I conjecture that, whether or not this is the case, there exists an equilibrium in which innovations are disclosed if and only if they exceed a common cutoff. Moreover, this equilibrium is more efficient than the old one.

In Chapter 4, my coauthor and me show that the set of all preferences over a totally-ranked set of alternatives is a complete lattice when ordered according to single-crossing dominance, and characterise joins and meets of arbitrary sets of preferences. We also give necessary and sufficient conditions on the underlying pre-order of the alternatives for the existence of joins and meets, as well as for their uniqueness. We present five applications. In particular

1. We derive monotone comparative statics results showing that (a) when the set of preferences of a group agents increases, so does the set of alternatives preferred by all agents, and (b) when a set known to contain the true preferences of an agent increases, so does the smallest set guaranteed to contain her preferred alternative(s).

2. We characterise a general class of 'maxmin' preferences as precisely minimum upper bounds with respect to 'more ambiguity-averse than'.

3. We characterise when aggregation of individual preferences (respecting a suitable Pareto criterion) is possible if, for some given pairs of alternatives, society may rank one above the other only with the consent of all individuals.

Chapter 1

Monotone Decision Rules and Supermodularity

Abstract

We study decision problems under uncertainty involving the choice of a rule mapping states into actions. We show that for any rule, there exists an increasing rule generating higher expected value for *all* payoff functions that are supermodular in action and state. We present applications to problems of taxation, betting, and price-discrimination in markets with demand externalities. We then consider rules mapping noisy signals of the state into actions. Under some conditions, optimal rules are increasing when (a) several agents commit to a shared rule or (b) the relationship between signal and state is ambiguous. Moreover, standard informativeness criteria apply.

1 Introduction

Many decision problems entail the choice of a decision rule; that is, a map from observables (or states) into actions. Often, the problem is separable across states. For instance, if F describes the distribution of the state $y \in \mathbb{R}$, a decision maker who earns payoff $u(x, y)$ from action $x \in \mathbb{R}^n$ when the state is y, picks a function $\psi : \mathbb{R}^n \to \mathbb{R}$ in order to maximise

$$\int u(\psi(y), y) \mathrm{d}F(y).$$ [1]

1. An optimal ψ exists if u is measurable, $u(\cdot, y)$ is bounded for all $y \in \mathbb{R}$, and $\sup_{x \in \mathbb{R}^n} u(x, \cdot)$ is F-integrable.

Under what conditions on u is it optimal to choose an increasing rule ψ? The main result of Milgrom and Shannon (1994) implies that if u has single-crossing differences in (x, y) then, under mild assumptions, for any rule ψ there exists an increasing rule ϕ such that $u(\phi(y), y) \geq u(\psi(y), y)$ for all $y \in Y$.[2] In particular,

$$\int_Y u(\phi(y), y) \mathrm{d}F(y) \geq \int_Y u(\psi(y), y) \mathrm{d}F(y). \tag{1}$$

In many settings, however, we require that (1) holds for *all* functions u within a given set $\mathcal{U}$. If this is the case, we say that ϕ $\mathcal{U}$-dominates ψ. Under what conditions on $\mathcal{U}$ can one find, for any rule ψ, an increasing rule ϕ that $\mathcal{U}$-dominates ψ? Our main result is that, if the decision maker is allowed to take a random action x after observing y, then this is the case if $\mathcal{U}$ is the set of all supermodular functions (denoted $\mathcal{U}_{\text{SPM}}$).[3] The rule ϕ is constructed by rearranging ψ without altering the ex-ante distribution of actions induced by F; we call ϕ the *increasing rearrangement* of ψ *with respect to* F.[4] Thus, even if the decision maker does not know the functional form of her payoff function u – merely that it is supermodular, knowing F enables her to improve on any non-increasing rule ψ. Conversely, if $\mathcal{U}$ is any set of payoffs such that, for any F, any rule ψ is $\mathcal{U}$-dominated by its increasing rearrangement with respect to F, then all elements of $\mathcal{U}$ must be supermodular.

The result that any decision rule is $\mathcal{U}_{\text{SPM}}$-dominated by its increasing rearrangement has many interesting applications. We present five problems in which relatively weak assumptions ensure that optimal rules are monotone.

Applications

1. A social planner wishes to implement a subsidy based on a household-specific measure of need for financial support, subject to a budget constraint. If the marginal utility of the transfer is increasing in need, the planner may restrict attention rules that transfer more resources to households who are

2. It is sufficient that $\arg\max_{x \in X} u(x, y)$ in non-empty and has a maximal element for all $y \in Y$.

3. A random rule ψ is increasing if the support of $\psi(y')$ lies above that of $\psi(y)$ whenever $y' > y$. If randomisation is not allowed, then any rule ψ is $\mathcal{U}_{\text{SPM}}$-dominated by an increasing rule ϕ if and only if F is a continuous distribution or a discrete uniform distribution.

4. If X is multi-dimensional, each of the components of ψ is rearranged in such a way that, in each state, the (random) action taken is *comonotone*; that is, any two elements of its support are ordered component-wise.

more in need. Importantly, the planner may improve on any other rule without knowing the quantitative aspects of the households' utility (which are hardly ever known in practice), if she knows the distribution of need in the population.

2. Two agents place monetary bets over a random outcome. If one of the agents is more optimistic about the outcome in a MLR-sense then, under any Pareto optimal bet, higher outcomes induce larger transfers to her. Moreover, this holds whether agents maximise expected utility or *cautious* expected utility.

3. A monopolist price-discriminates based on buyers' wealth levels in a market with a positive demand externality. If production costs are negligible, demand is normal and wealthier buyers are less price-sensitive, then a profit-maximising seller would charge them more. In particular, shifting the weight of high prices on wealthier buyers without altering the overall distribution of prices increases profits. The reason is twofold: first, since wealthier buyers are less price-sensitive, this boosts the demand externality; second, since the good is normal, this increases the average price per unit sold.

4. An insurer sells a single product to a group of buyers with varying risk-aversion who are subject to different shocks. If she can only condition insurance payments on a noisy signal of the buyers' shocks, she should offer an insurance that pays more when signals indicating larger shocks are observed. Moreover, if the insurer gains access to a signal that is more accurate in Lehmann's sense, she may achieve higher profits whilst making all agents better-off.

5. A firm has access to a noisy signal of its marginal cost, which is unobservable. Before observing the signal, the firm commits to a signal-contingent level of output. The firm is ambiguity-averse, and does not know the joint distribution of signal and marginal cost – it merely knows the marginal distribution of the signal. We show that the firm should produce more after observing signals that indicate lower costs, and derive a condition on the structure of the ambiguity enabling the firm to achieve higher ex-ante utility.

A number of authors studied other problems in economics using techniques similar to those we draw on here. Landsberger and Meilijson (1994) show that for any

non-comonotone allocation of a random endowment to a group of risk-averse agents, there exists a comonotone allocation raising the expected utility of all agents, regardless of their specific utility functions. Arieli and Babichenko (2016) consider a problem of Bayesian persuasion in which a single sender sends private messages to multiple receivers. They show that, if the state is binary, the same signalling policy is optimal for any sender with a payoff function that is supermodular in the receivers' actions. Moreover, the signals are increasing in the state. Meyer (2018) obtains a general result that applies to constrained problems in which the marginal distributions for each component of the (multi-dimensional) action are predetermined. She links the degree of heterogeneity among these marginals to the value of the (unique) increasing rule when payoffs are supermodular and symmetric. Her result applies to the setting studied by Arieli and Babichenko (2016); she shows that heterogeneity in the persuadability of receivers is harmful to the sender. In the context of statistical estimation, Chernozhukov, Fernandez-Val, and Galichon (2009) show that the monotone rearrangement of an estimate of a monotone function improves upon the original estimate whenever the latter is not monotone. Similarly, rearranging the upper and lower bound of a confidence band for monotone functions reduces its length and increases its coverage probability.

The rest of the paper is divided in two sections. Section 2 contains the theoretical results, which apply to a general decision environment. Applications $1-3$ are discussed in Sections 3.1 – 3.3, respectively. Sections 3.4 and 3.5 contain specific results about choice under uncertainty and informativeness, respectively. Applications 4 and 5 make up Examples 1 and 2 of both sections.

2 Theory

This section contains the main theoretical results of the paper, expressed in their most general form. A series of applications are presented in Section 3.

Consider a decision maker who takes an action x after observing the state of the world y. Suppose that y is distributed according to some (cumulative) distribution function $F : \mathbb{R} \to [0,1]$ with support $Y \subseteq \mathbb{R}$, and that, for some $n \in \mathbb{N}$, $x \in X =$

$X_1 \times \cdots \times X_n$ where $X_i \subset \mathbb{R}$ is compact for all $i \leq n$. We allow the decision maker to randomise over X after observing any signal $y \in Y$. A *(randomised) decision rule* is a collection $\psi = \{\psi(\cdot|y)\}_{y \in Y}$ of distribution functions $\psi(\cdot|y) : \mathbb{R}^n \to [0,1]$ with support within X, indexed by $y \in Y$. The value $\psi(x|y)$ is the probability that the decision maker takes an action lower or equal to x (in the component-wise order) after observing state y; ψ is *deterministic* if, for all $y \in Y$, $\psi(\cdot|y)$ is degenerate (that is, it assigns probability 1 to some action $x \in X$). In this case, denote by $\psi(y)$ the action taken after observing state y.[5] Let D_F be the set of decision rules, and let $\hat{D}_F \subset D_F$ be the set of deterministic decision rules.[6]

Definition 1. *A decision rule ψ is* increasing *if, for any $(x, y) \in \mathbb{R}^n \times Y$ the following holds. If $\psi(x|y) < 1$, for any $y' > y$, we obtain $\psi(x|y') = 0$.*

Intuitively, ψ is increasing if the support of $\psi(\cdot|y')$ lies above that of $\psi(\cdot|y)$ whenever $y' > y$. Note that, if ψ is deterministic, the definition reduces to the requirement that ψ is a (weakly) increasing function from Y to X.

An n-dimensional distribution G is *comonotone* if, for any two points $z, z' \in \mathbb{R}^n$ within its support, either $z \geq z'$ or $z' \geq z$. Equivalently, G is comonotone if and only if $G(z) = \min_{i \leq n} G_i(z_i)$ where G_i is the i^{th} one-dimensional marginal distribution of G for all $i \leq n$.[7] A decision rule $\psi \in D_F$ is comonotone if $\psi(\cdot|y)$ is a comonotone distribution for all $y \in Y$. Note that any deterministic decision rule is comonotone and, if $n = 1$, then any rule is comonotone. Throughout the discussion, we regard any two decision rules ψ and ψ' such that $\psi(\cdot|y) = \psi'(\cdot|y)$ for F-almost all y's as being essentially the same rule. In particular, if ψ is increasing (resp. comonotone), then any ψ' such that $\psi'(\cdot|y) = \psi(\cdot|y)$ for F-almost all y's is considered increasing (resp. comonotone) too.

Let $u(x, y)$ be the payoff to the decision maker if she takes action $x \in X$ and the state is $y \in Y$. Then, if she adopts a decision rule ψ, her expected payoff before

5. If Y is uncountable, we require that $\psi(x|\cdot)$ is a measurable function over Y (endowed with the Borel σ-algebra) for any $x \in X$. This implies that if ψ is deterministic, it may be described by a measurable map $\psi : Y \to X$. Throughout the discussion, we assume that all functions are measurable.

6. Formally, D_F is the quotient space of collections $\{\psi(\cdot|y)\}_{y \in Y}$ by the equivalence relation $\sim$ whereby $\psi \sim \psi'$ if and only if $\psi(\cdot|y) = \psi'(\cdot|y)$ for F-almost all y's; $\hat{D}_F$ is the set of equivalence classes in D_F that contain a deterministic rule.

7. See Dhaene et al. (2002) for a theoretical analysis of comonotonicity and a proof of this equivalence.

observing the state is

$$\int_Y \int_X u(x,y)\mathrm{d}\psi(x|y)\mathrm{d}F(y).^{8} \tag{2}$$

Definition 2. *For a set $\mathcal{U}$ of payoff functions and $\phi, \psi \in D_F$, ϕ* $\mathcal{U}$-dominates *ψ if*

$$\int_Y \int_X u(x,y)\,d\phi(x|y)\,dF(y) \geq \int_Y \int_X u(x,y)\,d\psi(x|y)\,dF(y) \tag{3}$$

for all $u \in \mathcal{U}$; ϕ strictly $\mathcal{U}$-dominates *ψ if (3) is strict for all $u \in \mathcal{U}$.*

Denote the i^{th} entry of a vector $z \in \mathbb{R}^m$ by z_i. Given $z, z' \in \mathbb{R}^m$, let $z \wedge z' = (\min\{z_1, z_1'\}, \ldots, \min\{z_m, z_m'\})$ and $z \vee z' = (\max\{z_1, z_1'\}, \ldots, \max\{z_m, z_m'\})$.

Definition 3. *A function $u : \mathbb{R}^m \to \mathbb{R}$ is* supermodular *if for any $z, z' \in \mathbb{R}^m$*

$$u(z \wedge z') + u(z \vee z') \geq u(z) + u(z'); \tag{4}$$

u is strictly supermodular *if (4) is strict unless $\{z \wedge z', z \vee z'\} = \{z, z'\}$; v is* (strictly) submodular *if $-v$ is (strictly) supermodular.*

Denote by $\mathcal{U}_{\text{SPM}}$ (resp. $\mathcal{U}_{\text{SSPM}}$) the set of (strictly) supermodular payoff functions. Given $\psi \in D_F$ and $i \leq n$, let $\psi_i = \{\psi_i(\cdot|y)\}_{y \in Y}$ be such that $\psi_i(\cdot|y)$ is the marginal distribution of $\psi(\cdot|y)$ over X_i. Intuitively, ψ_i is the 'projection' of ψ onto X_i.[9] Any rule $\psi \in D_F$ induces a distribution of action components $x_i \in X_i$ given by

$$M_{\psi_i F}(x_i) = \int_Y \psi_i(x_i|y)\mathrm{d}F(y). \tag{5}$$

Definition 4. *A* rearrangement of *a rule $\psi \in D_F$ with respect to F is a rule $\phi \in D_F$ such that $M_{\phi_i F} = M_{\psi_i F}$ for all $i \leq n$.*

The next result shows that any rule ψ admits a unique increasing and comonotone rearrangement with respect to F, and gives conditions on F and ψ ensuring that the rearrangement is deterministic. The distribution F is *continuous* whenever $F : \mathbb{R} \to [0,1]$ is a continuous function; F is a *discrete uniform* distribution if it puts mass $1/m$ on each of m points for some $m \geq 1$.

8. We restrict attention to functions $u : X \times Y \to \mathbb{R}$ such that $\int_Y \sup_{x \in X} |u(x,y)|\mathrm{d}F(y) < \infty$.

9. If ψ is deterministic (and, as a consequence, may be viewed as a map $\psi : Y \to X$) then ψ_i may be viewed as the function mapping $y \in Y$ to the i^{th} coordinate of $\psi(y)$.

Lemma 1. *For any $\psi \in D_F$, there exists a unique increasing and comonotone $\phi \in D_F$ that is a rearrangement of ψ with respect to F. Moreover, $\phi \in \hat{D}_F$ if (a) F is continuous, or (b) $\psi \in \hat{D}_F$ and F is a discrete uniform distribution. If $\phi \in \hat{D}_F$ then ϕ is the unique increasing rearrangement of ψ in $\hat{D}_F$.*

Proof of Lemma 1. We first show that ϕ exists. We start by defining the marginal distributions $\phi_i(\cdot|y)$ for all $i \leq n$ and $y \in Y$. Fix $y \in Y$. If F is continuous at y, then let

$$\phi_i(x_i|y) = \begin{cases} 0 & \text{if } M_{\psi_i F}(x_i) < F(y) \\ 1 & \text{if } M_{\psi_i F}(x_i) \geq F(y). \end{cases}$$

If instead F has an atom of mass $p > 0$ at y, let

$$\phi_i(x_i|y) = \begin{cases} 0 & \text{if } M_{\psi_i F}(x_i) < F(y) - p \\ \frac{1}{p}[M_{\psi_i F}(x_i) - (F(y) - p)] & \text{if } F(y) - p \leq M_{\psi_i F}(x_i) < F(y) \\ 1 & \text{if } F(y) \leq M_{\psi_i F}(x_i). \end{cases}$$

Since $M_{\psi_i F}(x_i)$ is a distribution function and X_i is compact, then $\phi_i(\cdot|y)$ is a well-defined distribution for all $y \in Y$.[10] Let $\phi \in D_F$ be given by

$$\phi(x|y) = \min_{i \leq n} \phi_i(x_i|y) \quad \text{for all} \quad x \in \mathbb{R}^n, y \in Y.$$

Then, ϕ is a well-defined decision rule, and it is comonotone.

To complete the proof of existence, it remains to show that ϕ is increasing and that it is a rearrangement of ψ. To show that ϕ is increasing, fix $y \in Y$ and $x \in \mathbb{R}^n$ such that $\phi(x|y) < 1$. Then $\phi_i(x_i|y) < 1$ for some $i \leq n$. This implies that $F(y) > M_{\psi_i F}(x_i)$. Then $\phi_i(x_i|y') = 0$ for any $y' > y$. If F does not have an atom at y', this holds since $F(y') \geq F(y) > M_{\psi_i F}(x_i)$. If F puts mass $p > 0$ on y' this holds since $M_{\psi_i F}(x_i) - F(y') \leq F(y) - F(y') \leq -p$. Hence $\phi(x|y') = 0$. Therefore ϕ is increasing.

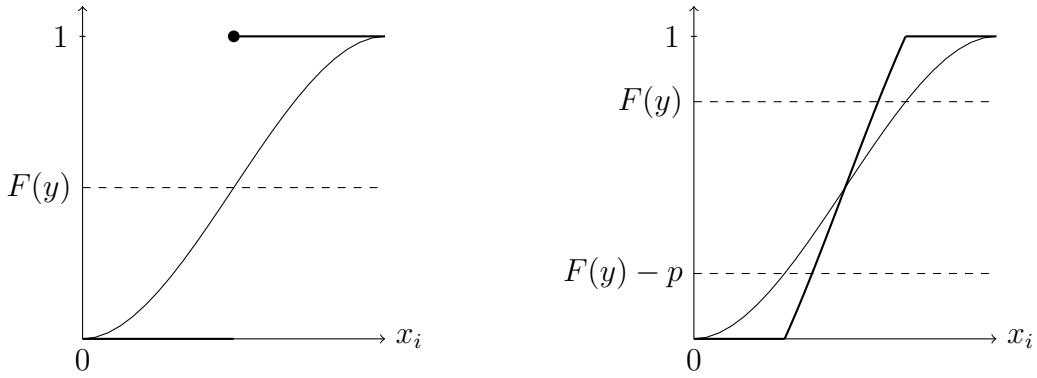
To show that ϕ is a rearrangement of ψ, fix $i \leq n$ and $x_i \in \mathbb{R}$. If $M_{\psi_i F}(x_i) = 0$, then $\phi_i(x_i|y) = 0$ for all $y \in Y$ such that $F(y) > 0$, so that $M_{\phi_i F}(x_i) =$

10. That is, $\phi_i(\cdot|y)$ is increasing, right-continuous and satisfies $\lim_{x_i \to -\infty} \phi_i(x_i|y) = 0$ and $\lim_{x_i \to \infty} \phi_i(x_i|y) = 1$.

$\int_Y \phi_i(x_i|y)\mathrm{d}F(y) = 0 = M_{\psi_i F}(x_i)$. Otherwise, let $y_0 = \inf\{y \in Y : M_{\psi_i F}(x_i) \leq F(y)\}$. Then, $\phi_i(x_i|y) = 1$ for $y < y_0$. Moreover, $\phi_i(x_i|y) = 0$ for any y such that $F(y) > F(y_0)$. Hence $\phi_i(x_i|y) = 0$ for F-almost all y's such that $y > y_0$. Let $p \geq 0$ be the probability mass that F puts on y_0. Then $M_{\psi_i F}(x_i) \in [F(y_0) - p, F(y_0)]$. Therefore $M_{\phi_i F}(x_i) = \int_Y \phi_i(x_i|y)\mathrm{d}F(y) = F(y_0) - p + p\phi_i(x_i|y_0) = M_{\psi_i F}(x_i)$. Since i and $x_i \in \mathbb{R}$ are arbitrary, ϕ is a rearrangement of ψ with respect to F. This proves existence. See the Appendix for a proof that ϕ is unique.

If F is continuous, it is clear from the definition of ϕ that $\phi \in \hat{D}_F$. Suppose that F is a discrete uniform distribution. Label $y_1 < \cdots < y_m$ the elements of Y. Then, to prove that $\phi \in \hat{D}_F$, it suffices to show that $\phi_i(\cdot|y_j)$ is degenerate for all $i \leq n$ and $j \leq m$. Fix i and j, and note that $\phi_i(x_i|y_j) = (1 + (mM_{\psi_i F}(x_i) - j) \wedge 0) \vee 0$. Moreover, since $\psi \in \hat{D}_F$, the image of $M_{\psi_i F}$ is within $\{0, 1/m, \ldots, (m-1)/m, 1\}$. It follows that $\phi_i(\cdot|y_j)$ is degenerate with value $\min\{x_i \in X_i : j/m \leq M_{\psi_i F}(x_i)\}$.

To prove the last part, note that any deterministic rule is comonotone. Therefore the uniqueness of the increasing rearrangement follows from the uniqueness of the increasing and comonotone rearrangement. □



(a) F is continuous at y (b) F has an atom of size p at y

Figure 1: $\phi_i(\cdot|y)$ (thick line) and $M_{\psi_i F}$ (thin line) as functions of x_i.

The construction of $\phi_i(\cdot|y)$ is illustrated in Figure 1. If F is continuous at y, then $\phi_i(\cdot|y)$ is degenerate and puts all mass on the smallest x_i such that $M_{\psi F}(x_i)$ lies above $F(y)$. If F has an atom of size $p > 0$ at y, then the support of $\phi_i(\cdot|y)$ lies within $[x_i', x_i'']$ where x_i' (resp. x_i'') is the smallest x_i such that $M_{\psi_i F}(x_i)$ lies above $F(y) - p$ (resp. $F(y)$). More specifically, $\phi_i(\cdot|y)$ is an affine transformation of $M_{\psi_i F}$ on $[x_i', x_i'']$.

Theorem 1. *Let $\psi \in D_F$ and let ϕ be the increasing and comonotone rearrangement of ψ with respect to F. Then*

1. *ϕ $\mathcal{U}_{SPM}$-dominates ψ*
2. *no other increasing and comonotone rule $\mathcal{U}_{SPM}$-dominates ψ*
3. *ϕ strictly $\mathcal{U}_{SSPM}$-dominates ψ unless ψ is increasing and comonotone.*

Theorem 1 is proved in the Appendix. It follows by combining Lemma 1 with the 'inequality for rearrangements' due to Lorentz (1953). Essentially, Lorentz's result states that Theorem 1 holds if F is the uniform distribution on $[0, 1]$ and D_F is replaced by $\hat{D}_F$. We generalise the result by allowing F to be arbitrary and the decision rules to be non-deterministic.

Theorem 1 allows to identify qualitative properties of the solutions to decision problems without the need to know their quantitative features. In Section 3, we give conditions on a range of decision problems ensuring that (a) they admit an optimal solution that is increasing or (b) all optimal solutions are increasing. Among other things, we show that a known informativeness criterion applies to decision problems in which it had not been used before (c.f. Section 3.5).

2.1 Limitations of deterministic rules

In general, restricting attention to deterministic decision rules entails a loss of generality. In particular, unless specific restrictions are imposed on F, one can find a non-inceasing deterministic rule ψ that is not $\mathcal{U}_{\text{SPM}}$-dominated by any increasing deterministic rule ϕ. In this case, Theorem 1 implies that the increasing and comonotone rearrangement of ψ with respect to F is not deterministic, even though ψ itself is. Counter-example 3 of Section 3.1 provides one such ψ, as well as a *randomised* increasing rule dominating it.

We showed in Lemma 1 that if F is continuous then the increasing (and comonotone) rearrangement of *any* rule ψ is deterministic.[11] We also showed that, if F is a discrete uniform distribution, the increasing (and comonotone) rearrangement

11. Hence, for the purposes of Theorem 1, restricting attention to deterministic rules *is* without loss of generality if F is continuous.

of any deterministic rule is deterministic. Hence, in these cases, Theorem 1 holds even if attention is restricted to deterministic rules.

Corollary 1 (Deterministic rules)**.** *Let F be a continuous distribution or a discrete uniform distribution. Let $\psi \in \hat{D}_F$ and let $\phi \in \hat{D}_F$ be the increasing rearrangement of ψ w.r.t. F. Then ϕ is the unique increasing deterministic rule that $\mathcal{U}_{SPM}$-dominates ψ. Moreover, ϕ strictly $\mathcal{U}_{SSPM}$-dominates ψ unless ψ is increasing.*

However, as the next result shows, this restriction on F is also *necessary* for any deterministic rule to admit a increasing rearrangement that is deterministic.

Proposition 1. *Suppose that X has more than one element and that, for any $\psi \in \hat{D}_F$, $\hat{D}_F$ contains an increasing rearrangement of ψ with respect to F. Then F is a continuous distribution or a discrete uniform distribution.*

Proposition 1 is proved in the Appendix. This result implies that, unless F is continuous or a discrete uniform distribution, one may find a non-increasing $\psi \in \hat{D}_F$ that is not $\mathcal{U}_{\text{SPM}}$-dominated by any increasing $\phi \in \hat{D}_F$.

Corollary 2. *Suppose that X has more than one element and that, for any $\psi \in \hat{D}_F$, there exists an increasing $\phi \in \hat{D}_F$ that $\mathcal{U}_{SPM}$-dominates ψ. Then F is a continuous distribution or a discrete uniform distribution.*

Proof. Fix $\psi \in \hat{D}_F$ and let $\phi \in \hat{D}_F$ be increasing and $\mathcal{U}_{\text{SPM}}$-dominate ψ. Since ϕ is deterministic, it is comonotone. Then Theorem 1 implies that ϕ is the increasing and comonotone rearrangement of ψ w.r.t. F. In particular, ϕ is a rearrangement of ψ w.r.t. F. Since ψ is arbitrary, result follows from Proposition 1. □

An analogue of Theorem 1 for deterministic rules holds if attention is restricted to payoff functions that are concave in the action. Let $\mathcal{U}_{\text{C}}$ be the set of payoff functions $u : X \times Y \to \mathbb{R}$ such that $u(\cdot, y)$ is concave for F-almost all $y \in Y$.

Corollary 3 (Concave payoffs)**.** *Given $\psi \in D_F$, there exists an increasing $\phi \in \hat{D}_F$ such that ϕ $\mathcal{U}_C \cap \mathcal{U}_{SPM}$-dominates ψ and, unless ψ is increasing and comonotone, ϕ strictly $\mathcal{U}_C \cap \mathcal{U}_{SSPM}$-dominates ψ.*

Proof. Fix $\psi \in D_F$ and let $\phi' \in D_F$ be the (possibly random) increasing and comonotone rearrangement of ψ with respect to F. Let $\phi \in \hat{D}_F$ be such that $\phi(y) = \int_X x \mathrm{d}\phi'(x|y)$ for all $y \in Y$. Then, Jensens' inequality implies that ϕ $\mathcal{U}_{\mathrm{C}}$-dominates ϕ'. Result follows from Theorem 1. □

2.2 The necessity of supermodularity

Theorem 1 implies that, for the purposes of simultaneous maximisation of supermodular payoffs, attention may be restricted to increasing (and comonotone) decision rules. From this perspective, it may be viewed as a result in monotone comparative statics. In this section, we show that Theorem 1 does not hold if supermodularity is weakened to the notion of single-crossing differences in (x, y), defined below. Moreover, we prove that supermodularity is tight in a specific sense.

Definition 5. *A function $u : \mathbb{R}^2 \to \mathbb{R}$ has* single-crossing differences in (x, y) *if for any $x, x', y, y' \in \mathbb{R}$ such that $x' > x$ and $y' > y$,*

$$u(x', y) \geq u(x, y) \Rightarrow u(x', y') \geq u(x, y')$$
$$u(x', y) > u(x, y) \Rightarrow u(x', y') > u(x, y').$$[12]

Suppose that $X \subseteq \mathbb{R}$ is compact and let $u : X \times Y \to \mathbb{R}$ be such that $u(\cdot, y)$ is continuous for all $y \in Y$. As noted in the introduction, Theorem 4 of Milgrom and Shannon (1994) implies that, if u has single-crossing differences in (x, y), then for any $\psi \in \hat{D}_F$ there exists an increasing $\phi \in \hat{D}_F$ such that $u(\phi(y), y) \geq u(\psi(y), y)$ for all $y \in Y$.[13] Therefore, (1) holds for any distribution F. Nevertheless, given payoffs u_1 and u_2 with single-crossing differences in (x, y), there may exist some non-increasing $\psi \in D_F$ that is not $\{u_1, u_2\}$-dominated by any increasing $\phi \in D_F$.

Counter-example 1. Let $Y \subset \mathbb{R}$ be bounded, $X = \mathbb{R}$, and consider $u_1(x, y) = -x^2/2$ and $u_2(x, y) = e^{-(x-y)^2/2}$. Note that u_1 is supermodular and u_2 has single-

12. Milgrom and Shannon (1994) introduced this concept as the 'single-crossing property'. We adopt the name used by Quah and Strulovici (2009).

13. Let ϕ be such that $\phi(y) = \sup D^*(y)$ for all $y \in Y$ where $D^*(y) = \arg\max_{x \in X} u(x, y)$. Since $u(\cdot, y)$ is continuous and X is compact, then $D^*(y)$ is non-empty and compact. Hence $\phi(y) \in D^*(y)$, so $u(\phi(y), y) \geq u(\psi(y), y)$. Since u has single-crossing differences in (x, y), Theorem 4 of Milgrom and Shannon (1994) implies that $D^*(y)$ is increasing in y in the strong set order. Therefore ϕ is increasing.

crossing differences in (x, y) but it is not supermodular.[14] Let F be any distribution with support Y, and let $D_F^* = \arg\max_{\varphi \in D_F} \int_Y U(\varphi(y), y) \mathrm{d}F(y)$ where $U(x, y) = u_1(x, y) + u_2(x, y)$. Note that, for each $y \in \mathbb{R}$, $U(\cdot, y)$ has a unique maximum.[15] Let $\psi : Y \to \mathbb{R}$ be such that $\psi(y) = \arg\max_{x \in \mathbb{R}} U(x, y)$ for all $y \in Y$. Since Y is bounded, it follows that $D_F^* = \{\psi\}$. Suppose that, for some F, ψ is $\{u_1, u_2\}$-dominated by some $\phi \in D_F$. Then $\int_Y U(\phi(y), y) \mathrm{d}F(y) \geq \int_Y U(\psi(y), y) \mathrm{d}F(y)$. Since $\psi \in D_F^*$, it follows that $\phi \in D_F^*$. Hence $\phi = \psi$. Therefore no rule $\{u_1, u_2\}$-dominates ψ other than itself. Moreover, one can show that ψ is decreasing on $\left(-\infty, -1 - e^{-1/2}\right]$, increasing on $\left[-1 - e^{-1/2}, 1 + e^{1/2}\right]$, and decreasing on $\left[1 + e^{1/2}, \infty\right)$.[16]

The next result shows that, unless specific assumptions are imposed on F, supermodularity is necessary for Theorem 1 to hold. The proof is in the Appendix.

Proposition 2. *Let $\mathcal{U}$ be an arbitrary set of payoff functions $u : X \times Y \to \mathbb{R}$. Suppose that, for any distribution F having support within Y, and any $\psi \in D_F$, some increasing rearrangement of ψ with respect to F $\mathcal{U}$-dominates ψ. Then $\mathcal{U} \subseteq \mathcal{U}_{SPM}$.*

3 Applications

To simplify exposition, we restrict attention to scalar deterministic decision rules throughtout this section. That is, we impose $n = 1$ so that $X \subset \mathbb{R}$, and constrain the decision maker to pick maps $\psi : Y \to X$ (i.e. elements of the set $\hat{D}_F$). In order to be able to apply our main result (c.f. Corollary 1), we assume that the distribution F is continuous. However, as noted in Section 2, the result holds for arbitrary distributions F if the decision maker is allowed to pick random decision

14. A smooth function $v : \mathbb{R}^2 \to \mathbb{R}$ is supermodular if and only if its cross-derivative is positive. In particular, the function $v(x, y) = -(x - y)^2$ is supermodular. Then u_2 is a increasing transformation of a supermodular function. Hence it has single-crossing differences in (x, y). However, u_2 is not supermodular since $\frac{\partial^2 u_2}{\partial x \partial y} = u_2(x, y)\left(1 - (x - y)^2\right)$ is positive if and only if $|x - y| \leq 1$.

15. That is because $U(\cdot, y)$ is strictly concave and $U(x, y) \to -\infty$ if $x \to \pm\infty$.

16. Since $U(\cdot, y)$ is strictly concave, $\psi(y)$ satisfies the first-order condition $\frac{\partial U}{\partial x}(\psi(y), y) = 0$. Total differentiation with respect to y shows that $\psi'(y)$ has the same sign as $\frac{\partial^2 u_2}{\partial x \partial y}(\psi(y), y)$. The first-order condition implies that $\psi'(y) = 0$ if and only if $y = \pm\left(1 + e^{-1/2}\right)$.

rules (c.f. Theorem 1). Accordingly, all the results proved in this section continue to hold for arbitrary distributions F if the decision maker may pick random rules.

3.1 Welfare Transfers

Consider a social planner wishing to implement a welfare transfer based on a household-specific measure $y \in Y$ of need for financial support. Let $v(x, y)$ be the utility of households with need $y \in Y$ who receive a transfer $x \in X$, where the function $v : X \times Y \to \mathbb{R}$ is increasing in x and supermodular. This means that households who receive larger transfers are better-off, and that the marginal utility of transfers is increasing in need.[17] Let F describe the distribution of need in the population. A *transfer* is a function $\psi \in \hat{D}_F$ whereby $\psi(y)$ is transferred to households with need $y \in Y$. A planner with a budget $b \in \mathbb{R}$ solves

$$\max_{\psi \in D_F} \int_Y v(\psi(y), y) \mathrm{d}F(y) \quad \text{subject to} \quad \int_Y \psi(y) \mathrm{d}F(y) \leq b.$$

This can be viewed as an optimal taxation problem, similar to those studied by Ramsey (1927) and Mirrlees (1971).[18] Note that $\int_Y v(\psi(y), y) \mathrm{d}F(y)$ is the average *welfare* generated by the transfer, whereas $\int_Y \psi(y) \mathrm{d}F(y)$ is its *cost*. To ensure that the problem has a solution, suppose that Y is compact and that v is continuous.

The following result states that any non-increasing transfer can be rearranged into an increasing transfer that produces higher welfare, without altering its cost. Crucially, this rearrangement is not sensitive to the functional form of the households' utility u, as long as u is supermodular. Thus, a planner facing a population of households with supermodular utility may restrict attention to increasing transfers even if, as it almost always is the case in practice, she does not know the quantitative aspects of the households' utility function.

Proposition 3. *For any non-increasing $\psi \in \hat{D}_F$, the increasing rearrangement of ψ with respect to F has same cost and generates higher welfare for any supermodular*

17. For example, pick $v(x, y) = g(x - y)$ for some increasing and concave function $g : [0, \infty) \to \mathbb{R}$ and view y as the negative of wealth. Alternatively, y can be a poverty index that is not based on wealth alone (such as the Global Multi-dimensional Poverty Index).

18. Note that the households have no outside option and they are assumed to report y truthfully regardless of incentive compatibility.

v. If v is strictly *supermodular, then all optimal transfers are increasing.*

Proof. Fix $\psi \in \hat{D}_F$ and let ϕ be its increasing rearrangement with respect to F. Since F is continuous, Corollary 1 implies that ϕ $\mathcal{U}_{\text{SPM}}$-dominates ψ. Since $v \in \mathcal{U}_{\text{SPM}}$, ϕ produces higher welfare than ψ. Since ϕ is a rearrangement of ψ then ϕ and ψ induce the same *distribution* of transfers x. In particular, $\int_Y \phi(y)\mathrm{d}F(y) = \int_Y \psi(y)\mathrm{d}F(y)$; that is, ϕ and ψ have the same cost. To prove the second part, suppose that ψ is not increasing. Then Corollary 1 implies that ϕ strictly $\mathcal{U}_{\text{SSPM}}$-dominates ψ. Since $v \in \mathcal{U}_{\text{SSPM}}$, then ϕ induces strictly higher welfare than ψ. Hence ψ can't be optimal. □

The following examples show that the (unique) optimal transfer may be *decreasing* if v is not supermodular, or if F is not continuous. However, in the latter case, this only holds under the constraint of deterministic transfers.

Counter-example 2: *if v is not supermodular.* Let $X = [0, 1]$, and let $v(x, y) = (x - y)^2$ where $y \leq 0$ denotes the negative of wealth. This means that the marginal revenue to transfers is *increasing* in households' wealth. Note that v has single-crossing differences in (x, y) since it is strictly increasing in x. However, it is strictly *sub*modular. It is easy to show that in this case the optimal rule takes the form $\psi(y) = \mathbb{I}\{F(y) \leq b\}$. That is, the planner should implement the largest feasible transfers and target all households that are sufficiently wealthy. In particular, the optimal transfer is *decreasing*.[19]

Counter-example 3: *if F is not continuous.* Let $Y = \{1, 2\}$ where 1 (resp. 2) denotes low (resp. high) need and assume that a proportion $\rho \in (1/2, 1)$ of the households is in high need.[20] Suppose that $u(x, y) = x^2 y$ so that the marginal return to transfers is increasing in their size. It is easy to show that, for any $b > 0$, the planner should transfer all resources to the population in low need if $\rho > 2/3$. By targeting the minority, the planner is able to transfer a higher amount to each household. Since $u(\cdot, 1)$ is sufficiently convex, this is optimal.[21]

19. More generally, if v is submodular, then rearranging *any* transfer into a decreasing transfer increases welfare without altering its cost; if v is strictly submodular, then all optimal transfers are decreasing. This can be seen by adapting the argument in the proof of Proposition 3.

20. If $\rho \in \{1/2, 1\}$, then F is a discrete uniform distribution and Proposition 3 holds.

21. Corollary 3 implies that if $u(\cdot, y)$ is concave for all $y \in Y$ then there exists an optimal increasing transfer, whether or not F is continuous.

If the planner is not constrained to transfer the same amount of resources to households with the same level of need, then an increasing transfer is optimal.[22] In particular, if $X = [0, x_0]$, then it is optimal to transfer x_0 to a fraction $b/(x_0\rho)$ of the households in high need, and to transfer nothing to the rest of the population. This enables the planner to achieve strictly higher welfare than if she were constrained to use a deterministic transfer. However, *horizontal equity* is violated since, among the households in high need, some receive higher transfers than others.[23]

3.2 Bets with subjective beliefs

We analyse a setting in which supermodular payoffs arise naturally from the combination of their monotonicity and the monotonicity of the likelihood ratio of events.

Consider two agents who agree on a bet over an outcome $y \in Y$. A *bet* is a function $\psi : Y \to X$ whereby $\psi(y)$ is transferred from Agent 2 to Agent 1 if the outcome is y. For $i = 1, 2$, let $u_i(x)$ be the utility to Agent i if the net transfer is $x \in X$, where the function $u_1 : X \to \mathbb{R}$ is increasing in x and $u_2 : X \to \mathbb{R}$ is decreasing. For $i = 1, 2$, Agent i believes that the outcome y has distribution F_i, which admits a density f_i with support Y. Then, the value to Agent i of a bet ψ is

$$U_i(\psi) = \int_Y u_i[\psi(y)]\mathrm{d}F_i(y).$$

Suppose that Agent 1 is more optimistic about the outcome in the sense that F_1 is a MLR-shift of F_2. The following result states that the Pareto optimal expected utility pairs can be reached by transferring more to Agent 1 when the outcome is higher. Moreover, if F_1 is a strict MLR-shift of F_2 and u_1 is strictly increasing, any Pareto optimal bet must be of this form.[24]

Proposition 4. *Any Pareto optimal expected utility pair can be achieved by an increasing bet. If F_1 is a* strict *MLR-shift of F_2 and u_1 is strictly increasing, then*

22. A transfer of this kind is an example of a randomised decision rule. They are defined in Section 2. In particular, see Definition 1 of increasing randomised decision rules.

23. Stiglitz (1982) argues that welfare maximisation and even Pareto optimality are sometimes at odds with the principle of horizontal equity, particularly in the context of indirect taxation. He describes economically relevant settings in which payoffs are not concave and random taxation is optimal.

24. F_1 is a strict MLR-shift of F_2 if $f_1(y)/f_2(y)$ is strictly increasing on Y.

all Pareto optimal bets are increasing.

Proof. Fix a bet ψ and let ϕ be the increasing rearrangement of ψ with respect to F_2. Then ϕ and ψ induce the same distribution of transfers under F_2, so $U_2(\phi) = U_2(\psi)$. Let $u(x, y) = u_1(x)f_1(y)/f_2(y)$. Since u_1 is increasing and F_1 is a MLR-shift of F_2, then u is supermodular. From Corollary 1, if $F = F_2$ then ϕ $\mathcal{U}_{\text{SPM}}$-dominates ψ. Moreover, for any transfer φ, $\int_Y u(\varphi(y), y)\mathrm{d}F_2(y) = U_1(\varphi)$. Therefore $U_1(\phi) \geq U_1(\psi)$. Hence, if ψ is Pareto optimal, then so is ϕ. This proves the first part. To prove the last part, note that if u_1 is strictly increasing and F_1 is a strict MLR-shift of F_2, then u is strictly supermodular. Suppose that ψ is not increasing. Then Corollary 1 implies that, if $F = F_2$, ϕ strictly $\mathcal{U}_{\text{SSPM}}$-dominates ψ. From above, it follows that $U_1(\phi) > U_1(\psi)$. Hence ψ is not Pareto optimal. □

Proposition 4 can be extended to the case where agents maximise *cautious expected utility* (c.f. Cerreia-Vioglio, Dillenberger, and Ortoleva 2015). These preferences satisfy a weaker form of the independence axiom which captures a leaning towards certain prospects. If agents are cautious expected utility maximisers, the value of a bet ψ to Agent i is

$$C_i(\psi) = \inf_{u \in \mathcal{U}_i} u^{-1}\left[\int_Y u[\psi(y)]\mathrm{d}F_i(y)\right]$$

where $\mathcal{U}_1$ (resp. $\mathcal{U}_2$) is a compact set of strictly increasing (resp. decreasing) utility functions. Note that these preferences reduce to expected utility whenever U_i's are singletons. We showed in the proof of Proposition 4 that, under expected utility, rearranging any bet ψ with respect to F_2 into an increasing bet ϕ induces a Pareto improvement. Furthermore, Since the construction of ϕ is not sensitive to the particular choice of u, we are able to deduce that

$$u^{-1}\left[\int_Y u[\phi(y)]\mathrm{d}F_i(y)\right] > u^{-1}\left[\int_Y u[\psi(y)]\mathrm{d}F_i(y)\right] \tag{6}$$

holds for *all* $u \in \mathcal{U}_i$ and $i = 1, 2$. Hence, Proposition 4 extends to the cautious expected utility case.[25]

25. In fact, Proposition 4 holds for the larger class of *convex* preferences, introduced by Cerreia-Vioglio

Proposition 5. *Any Pareto optimal cautious expected utility pair can be achieved by an increasing bet. If F_1 is a* strict *MLR-shift of F_2, then all Pareto optimal bets are increasing.*

Proof. Let ψ be a bet and let ϕ be its increasing rearrangement with respect to F_2. The proof of Proposition 4 implies that $U_2(\phi) = U_2(\psi)$ for *any* function u_2. If u_2 is strictly decreasing, $U_2(\phi) = U_2(\psi)$ is equivalent to $u_2^{-1}[U_2(\phi)] = u_2^{-1}[U_2(\psi)]$. Since elements of $\mathcal{U}_2$ are strictly decreasing, $C_2(\phi) = C_2(\psi)$. Similarly, $U_1(\phi) \geq U_1(\psi)$ for any increasing function u_1. If u_1 is strictly increasing then $U_1(\phi) \geq U_1(\psi)$ is equivalent to $u_1^{-1}[U_1(\phi)] \geq u_1^{-1}[U_1(\psi)]$. Since elements of $\mathcal{U}_1$ are strictly increasing, $C_1(\phi) \geq C_1(\psi)$. Thus, if ψ is Pareto optimal, then so is ϕ. This proves the first part. The second part follows by a similar reasoning. □

3.3 Price discrimination and demand externality

Consider a market for a single good with a continuum of buyers and a single seller. Let $q(x, y, z)$ be the demand of a buyer with wealth y facing price x when the average demand for the good is z. Suppose that q is increasing in z. Thus, the good exhibits a positive demand externality.[26] Assume also that q is increasing in wealth y, so that the good is normal, and that q is supermodular in prices and wealth (x, y). That is, for any $x' < x''$ and any z, $q(x', y, z) - q(x'', y, z)$ is decreasing in y. Hence the demand of wealthier buyers is less sensitive to changes in prices.

As an example, consider the demand for a particular software package or operating system. The positive externality is due to network effects and wealthier buyers, such as large firms, have higher demand (than small firms or individuals, say) as they have more employees and therefore require more licenses. Large firms are also less price-sensitive as their expenditure on the good is a smaller share of their total costs. The variable y may also be interpreted as the buyers' *brand preference* for the company selling the good. Buyers with a stronger preference demand more of

(2009). The value of ψ under these preferences may be represented as $\inf_{u \in U_i} V_i(\int_Y u[\psi(y)] \mathrm{d}F_i(y), u)$ for a set U_i of increasing functions u and a family $\{V_i(\cdot, u)\}_{u \in U_i}$ of increasing functions. This representation reduces to cautious expected utility if the elements of U_i are strictly increasing and $V_i(\cdot, u) = u^{-1}$ for all $u \in U_i$. The argument used in the proof of Proposition 5 extends naturally to this case.

26. The externality is *global*, as it depends on average demand. Farrell and Saloner (1985) and Katz and Shapiro (1985) study models of global demand externalities, but do not allow price discrimination.

the good, all other things being equal, and are less sensitive to changes in its price.

Suppose that the seller is able to price discriminate based on individual wealth.[27] In the market for software packages, the fact that firms offer a range of vertically differentiated products which all share the same (small) marginal cost of production may be viewed as an attempt to price discriminate. Let the distribution of wealth in the population be described by F. A *pricing rule* for the seller is a function $\psi \in D_F$ such that buyers with wealth $y \in Y$ are charged $\psi(y) \in X$.

As noted at the start of Section 3, our main result applies if F is continuous or if the seller is able to implement a random pricing rule. The former assumption requires the seller to have sufficient information about buyers' wealth to be able to price discriminate finely among them. The latter requires the seller to be able to charge different prices to buyers with the same wealth (using e.g. lotteries for discounts). Given a pricing rule ψ, average demand, denoted by z_ψ, solves

$$z_\psi = \int_Y q(\psi(y), y, z_\psi) \mathrm{d}F(y). \tag{7}$$

To ensure that (7) has a unique solution for any ψ, suppose that Y is compact, that q is strictly positive and that, for any (x, y), $q(x, y, \cdot)$ is concave and $q(x, y, z) < z$ for z large enough.[28] Suppose that the seller can produce at no cost. Then her profit-maximisation problem is

$$\max_{\psi \in D_F} \int_Y \psi(y) q(\psi(y), y, z_\psi) \mathrm{d}F(y). \tag{8}$$

To ensure that the problem has a solution, suppose further that q is continuous. The following result states that it is optimal for the seller to charge wealthier buyers (who are also less price-sensitive) more.

27. There is a network literature studying price discrimination by a single seller in a market with a positive demand externality; however, the externality is local and arises from a network of social interactions. In the models of Candogan, Bimpikis, and Ozdaglar (2012) and Bloch and Quérou (2013) as well as in ours, the monopolist has full information. Sundararajan (2007) and Fainmesser and Galeotti (2015) limit the information that the seller has about the network. Using the results of Section 3.4, our framework can be extended to the case where the seller only observes a noisy signal of individual wealth.

28. Note that, since q is decreasing in x and increasing in y, the function $g_\psi(z) = \int_Y q(\psi(y), y, z) \mathrm{d}F(y)$ has image in $[0, z_0]$ for any z_0 such that $q(\min X, \max Y, z_0) < z_0$. To prove existence, apply Tarski's theorem to g_ψ on $[0, z_0]$. To prove uniqueness, note that g_ψ is concave and $g_\psi(0) > 0$. Moreover, for any $z' \in (0, z_\psi)$, $z' = (z'/z_\psi) g_\psi(z_\psi) < (z'/z_\psi) g_\psi(z_\psi) + (1 - z'/z_\psi) g_\psi(0) \leq g_\psi(z')$, where the last inequality follows from the concavity of g_ψ. Hence z' is not a fixed-point of g_ψ. The same holds if $z' > z_\psi$.

Proposition 6. *For any pricing rule ψ, the increasing rearrangement of ψ with respect to F induces higher revenue. If $q(x, y, z)$ is strictly supermodular in (x, y) for all $z \geq 0$, then the optimal pricing rules are increasing.*

Proof. Fix $\psi \in D_F$. Since F is a continuous and q is supermodular in (x, y) for all $z \geq 0$, Corollary 1 implies that the increasing rearrangement of ψ with respect to F (ϕ, say) $\mathcal{U}$-dominates ψ where $\mathcal{U} = \{q(\cdot, z) : z \geq 0\}$. This implies that $z_\phi \geq z_\psi$.[29] Let $v(x, y) = xq(x, y, z_\psi)$. Since $q(x, y, z_\psi)$ is increasing in y and supermodular in (x, y), then v is supermodular. Hence Corollary 1 implies that

$$\int_Y \phi(y)q(\phi(y), y, z_\psi)\mathrm{d}F(y) \geq \int_Y \psi(y)q(\psi(y), y, z_\psi)\mathrm{d}F(y). \tag{9}$$

Then, since q is increasing in z and $z_\phi \geq z_\psi$, $xq(x, y, z_\phi) \geq xq(x, y, z_\psi)$ for all (x, y). Hence

$$\int_Y \phi(y)q(\phi(y), y, z_\phi)\mathrm{d}F(y) \geq \int_Y \phi(y)q(\phi(y), y, z_\psi)\mathrm{d}F(y). \tag{10}$$

Combining (9) and (10) implies that ϕ induces weakly higher revenue than ψ. To prove the last part, suppose that ψ is not increasing. If $q(x, y, z)$ is strictly supermodular in (x, y) for all $z \geq 0$, then v is strictly supermodular. Then, Corollary 1 implies that (9) is strict. The previous reasoning shows that ψ is not optimal. □

Intuitively, rearranging the pricing rule into an increasing rule shifts the weight of high prices onto wealthier consumers without altering the overall distribution of prices. Since revenue is supermodular in prices and wealth, other things being equal, this increases revenue. However, the shift also affects average demand. Indeed, since demand is supermdoular in prices and wealth, it increases too. This further increases revenue. For this reason, the revenue-maximisation problem has a monotone solution even though it involves a fixed-point.

29. Note that, since ϕ $\mathcal{U}$-dominates ψ, $g_\phi(z) \geq g_\psi(z)$ for all $z \in [0, z_0]$ (c.f. footnote 28). Since z_ψ is the unique fixed-point of g_ψ on $[0, z_0]$, $z_\psi = \inf\{z \in [0, z_0] | z \geq g_\psi(z)\}$. Then $z < g_\psi(z) \leq g_\phi(z)$ for all $z < z_\psi$. Since z_ϕ is the unique fixed-point of g_ϕ on $[0, z_0]$, $z_\phi = \inf\{z \in [0, 1] | z \geq g_\phi(z)\}$. Hence $z_\phi \geq z_\psi$.

3.4 Choice under uncertainty

Consider a decision maker who has to take an action $x \in X$ before observing the state of the world $s \in S$, where $S \subseteq \mathbb{R}$. Let $u(x, s)$ be her utility if she takes action x and the state is s. Suppose that she has access to a signal y of the state with marginal distribution F and support Y. Let $P(\cdot|y)$ be the posterior distribution of the state s conditional on observing the signal realisation $y \in Y$. The family $P = \{P(\cdot|y)\}_{y \in Y}$ is (strictly) *FOSD-ordered* if for any $y'' > y'$, $P(\cdot|y'')$ is a (strict) FOSD-shift of $P(\cdot|y')$.[30] Intuitively, this means that the decision maker expects the state to be higher when she observes higher signals. A *decision rule* is a function $\psi \in D_F$ whereby the agent takes action $\psi(y)$ when she observes the signal y. The (ex-ante) *value* of ψ is

$$\int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y). \tag{11}$$

Given $\psi, \phi \in D_F$, a set $\mathcal{U}$ of payoff functions u and a set $\mathcal{P}$ of families P of posteriors, ϕ *$\mathcal{U} \times \mathcal{P}$-dominates* ψ if

$$\int_Y \int_S u(\phi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y) \geq \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y) \tag{12}$$

for all $(u, P) \in \mathcal{U} \times \mathcal{P}$. Moreover, ϕ *strictly* $\mathcal{U} \times \mathcal{P}$-dominates ψ if (12) is strict for all $(u, P) \in \mathcal{U} \times \mathcal{P}$.

Several known monotonicity conditions on u and P guarantee the existence of an increasing decision rule that maximises ex-ante value.[31] However, the optimal rule generally depends on the choice of u and P. The following result shows that rearranging a decision rule with respect to F into an increasing rule increases ex-ante value for *all* supermodular payoff functions and *all* FOSD-ordered families of

30. $P(\cdot|y'')$ is a strict FOSD-shift of $P(\cdot|y')$ if $P(s|y'') \leq P(s|y')$ for all $s \in S$ and the inequality is strict for some $s \in S$.

31. Athey and Levin (1998) derive a joint condition on u and P that is satisfied both if u is supermodular and P is FOSD-ordered, and if u has single-crossing differences in (x, s) and P is MLR-ordered (that is, $P(\cdot|y'')$ is a MLR-shift of $P(\cdot|y')$ whenever $y'' > y'$). From Karlin and Rubin (1956), it suffices that $\{u(\cdot, y)\}_{y \in Y}$ is a quasiconcave family with increasing peaks (QCIP) and P is MLR-ordered. From Quah and Strulovici (2009), it suffices that P is MLR-ordered and u has the interval-dominance order property, a weaker notion than both single-crossing differences in (x, s) and QCIP.

posteriors simultaneously.[32] Let $\mathcal{U}_{\text{SPM}}$ (resp. $\mathcal{U}_{\text{SSPM}}$) be the set of all (strictly) supermodular payoff functions, and let $\mathcal{P}_{\text{FOSD}}$ (resp. $\mathcal{P}_{\text{SFOSD}}$) be the set of all (strictly) FOSD-ordered families of posteriors.

Proposition 7. *Let $\psi \in D_F$ and let ϕ be its increasing rearrangement with respect to F. Then ϕ $\mathcal{U}_{SPM} \times \mathcal{P}_{FOSD}$-dominates ψ. Moreover ϕ strictly $\mathcal{U}_{SSPM} \times \mathcal{P}_{SFOSD}$-dominates ψ unless ψ is increasing.*

Proof. Let u be a (strictly) supermodular payoff function and let P be a (strictly) FOSD-ordered family. Then, the interim expected utility $v_{u,P}(x,y) = \int_S u(x,s)\mathrm{d}P(s|y)$ after observing signal y and taking action x is a (strictly) supermodular function of (x,y). Since F is continuous, Corollary 1 implies that ϕ $\mathcal{U}$-dominates ψ where $\mathcal{U} = \{v_{u,P} : u \in \mathcal{U}_{\text{SPM}}, P \in \mathcal{P}_{\text{FOSD}}\}$. Moreover, unless ψ is increasing, ϕ $\mathcal{U}'$-dominates ψ where $\mathcal{U}' = \{v_{u,P} : u \in \mathcal{U}_{\text{SSPM}}, P \in \mathcal{P}_{\text{SFOSD}}\}$. Result follows since $\int_Y \int_S u(\varphi(y),s)\mathrm{d}P(s|y)\mathrm{d}F(y) = \int_Y v_{u,P}(\varphi(y),y)\mathrm{d}F(y)$ for any $\varphi \in D_F$, payoff function u and family of posteriors P. □

Example 1 – Insurance for heterogeneous agents. A group of agents, indexed by i, are subject to shocks s_i. A seller offers insurance payments conditioned on noisy signals y_i of s_i, where y_i's have identical marginal distribution F.[33] An *insurance* contract is a rule $\psi \in D_F$ whereby agent i receives a net transfer $\psi(y_i)$ if the insurer observes $y_i \in Y$.[34] Let $v_i(x_i - s_i)$ be the utility of Agent i if she buys the insurance, the value of the shock is s_i, and the insurance transfer is x_i, where v_i is increasing and concave. Let $\bar{v}_i$ be the utility of Agent i if she does not buy the insurance. Suppose that the insurer sells a single product and wants to maximise profits subject to all agents purchasing it. Then, she solves

$$\min_{\psi \in D_F} \int_Y \psi(y)\mathrm{d}F(y) \quad \text{s.t.} \quad \int_Y \int_S v_i(\psi(y_i) - s_i)\mathrm{d}P_i(s_i|y_i)\mathrm{d}F(y_i) \geq \bar{v}_i \quad \forall i$$

32. The complete class theorems of Karlin and Rubin (1956) and Quah and Strulovici (2009) give conditions on payoffs $u(x,s)$ and experiments $G = \{G(s|y)\}_{s\in S}$ ensuring that, for any ψ, there exists an increasing ϕ such that $\int_Y u(\phi(y),s)\mathrm{d}G(y|s) \geq \int_Y u(\psi(y),s)\mathrm{d}G(y|s)$ for all $s \in S$ simultanesouly. However, ϕ depends on the choice of u and G.

33. The fact that y_i is contractible whereas s_i is not may be due to legal restrictions or to the fact that the insurer does not observe s_i. The fact that y_i's have identical marginal distribution is justified if y_i is determined by the risk profiles of agent i and the insurer is able to offer different products to agents with different risk profiles.

34. The net transfer may be decomposed as $t(y_i) - p$ where $t(y_i) \geq 0$ is the insurance payment and $p \geq 0$ is the price of insurance.

where P_i is the FOSD-ordered family of distributions of s_i conditional on y_i.[35]

I claim that it is optimal to offer an insurance that pays more whenever the observed signal is higher. Formally, there exists an optimal ψ that is increasing. Given any ψ, let ϕ be the increasing rearrangement of ψ with respect to F. For all i, since v_i is concave, then $u_i(x_i, s_i) = v_i(x_i - s_i)$ is supermodular. Then Proposition 7 implies that ϕ induces higher ex-ante utility than ψ to all agents. Thus, any agent purchasing ψ also buys ϕ. Moreover, as ϕ is a rearrangement of ψ, $\int_Y \phi(y)\mathrm{d}F(y) = \int_Y \psi(y)\mathrm{d}F(y)$. Hence, if ψ is optimal, then so is ϕ. This proves the claim. A similar reasoning shows that, if P_i's are strictly FOSD-ordered and v_i's are strictly concave for all i, then any optimal insurance must be increasing.

Example 2 – Ambiguity over the family of posteriors. Consider a firm facing uncertainty about its marginal cost, parametrised by s. Suppose that a noisy signal y of s is available and that the firm commits to a decision rule ψ whereby it supplies $\psi(y)$ units if the signal realisation is y. Let $u(x, s)$ be the firm's profits when producing x units. Suppose that marginal cost is decreasing in s, so that u is supermodular.[36] If the firm knew the marginal distribution F of y and the family P of posteriors of s given y, its profit-maximisation problem would be

$$\max_{\psi \in D_F} \int_Y \int_S u(\psi(y), s)\mathrm{d}P(s|y)\mathrm{d}F(y).$$

Suppose instead that the firm knows F but does not know P.[37] Moreover, it is nevertheless confident that P belongs to some set $\mathcal{P}$ of FOSD-ordered posterior families.[38] Following the theory of Gilboa and Schmeidler (1989), the firm is a *ambiguity-averse* if her ex-ante utility from $\psi \in D_F$ is $\inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), s)\mathrm{d}P(s|y)\mathrm{d}F(y)$.

35. If the insurer were able to offer different insurance contracts to each agent, then the problem would be separable across i's and could be solved using standard techniques.

36. This holds if $u(x, s) = p(x)x - c(x, s)$ where $p(x)$ is the inverse demand and c is the cost function.

37. For instance, suppose that the firm has access to a large number of i.i.d. realisations of y but only observed few i.i.d. realisations of (y, s), e.g. because it has only been active for a short period of time. Then, the firm can estimate F with confidence but does not have a trustworthy estimate of P.

38. For instance, the firm might conjecture that, for any $y \in Y$, $s = \alpha + \beta y + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$ and pick $\mathcal{P} = \{N(\alpha + \beta y, \sigma^2) : (\alpha, \beta, \sigma^2) \in A\}$ where $A \subset \mathbb{R}^2 \times \mathbb{R}_+$ is a confidence interval for $(\alpha, \beta, \sigma^2)$. In this case, the elements of $\mathcal{P}$ are FOSD-ordered if and only if $\beta \geq 0$ for any $(\alpha, \beta, \sigma^2) \in A$.

Then, the firm solves

$$\max_{\psi \in D_F} \inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y). \tag{13}$$

I claim that it is optimal to supply more after observing higher realisations of y (which signals lower marginal costs). Formally, there exists an optimal rule that is increasing. Given any ψ, let ϕ be its increasing rearrangement with respect to F. Since u is supermodular and elements of $\mathcal{P}$ are FOSD-ordered, Proposition 7 implies ϕ $\{u\} \times \mathcal{P}$-dominates ψ. Hence

$$\inf_{P \in \mathcal{P}} \int_Y \int_S u(\phi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y) \geq \inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), s) \mathrm{d}P(s|y) \mathrm{d}F(y).$$

Thus, if ψ is optimal, then so is ϕ.

Note that, in the absence of commitment, the firm may not be dynamically consistent in this setting. That is, a decision rule ψ may be optimal before observing the signal but $\psi(y)$ may not be an optimal action after observing the signal realisation y. Nevertheless, an ambiguity-averse firm who picks supply x *after* observing the signal y in order to maximise *interim* expected profit $v(x, y) = \inf_{P \in \mathcal{P}} \int_S u(x, s) \mathrm{d}P(s|y)$ should also adopt an increasing rule. Indeed, a result from Dziewulski and Quah (2016) implies that, for any rule ψ, there exists an increasing rule ϕ such that $v(\phi(y), y) \geq v(\psi(y), y)$ for any $y \in Y$.[39] However, as noted above, ϕ need not improve on ψ from an ex-ante perspective.[40]

Example 3 – Knightian uncertainty. Proposition 7 can also be applied to problems of choice under uncertainty with incomplete preferences. Consider a decision maker with utility function u who does not know the joint distribution J of signal y and state s. Suppose that she knows that J belongs to some (closed, convex) set $\mathcal{J}$. Following Bewley (2002), the decision maker has *Knightian preferences* if, for

39. Dziewulski and Quah (2016) show that if u is supermodular and, for any $y \in \mathbb{R}$, $(\Lambda(y), \geq)$ is a set of one-dimensional distributions endowed with the FOSD-order such that $\Lambda(y'') \geq \Lambda(y')$ in the strong set order whenever $y'' > y'$, then $w(x, y) = \inf_{\lambda \in \Lambda(y)} \int_{\mathbb{R}} u(x, s) d\lambda(s)$ is supermodular. Note that $v = w$ if $\Lambda(y) = \{P(\cdot|y) : P \in \mathcal{P}\}$.

40. In fact, there is more than one plausible way for the firm to update its beliefs about s after observing y. See Epstein and Le Breton (1993), Gilboa and Schmeidler (1993), Epstein and Schneider (2003) and Hanany and Klibanoff (2007) for a discussion.

any pair of decision rules ψ, ϕ she prefers ϕ to ψ if and only if

$$\int_{Y\times S} u(\phi(y), s)\mathrm{d}J(y,s) \geq \int_{Y\times S} u(\psi(y), s)\mathrm{d}J(y,s) \quad \forall J \in \mathcal{J}.$$[41]

Note that these preferences are incomplete over the set of all decision rules. Suppose that u is supermodular and that elements of $\mathcal{J}$ share the same marginal distribution F over Y and induce FOSD-ordered posteriors over S. Then Proposition 7 implies that, for any $\psi \in D_F$, the decision maker has a preference for the increasing rearrangement of ψ with respect to F over ψ itself. Therefore, an agent with a supermodular utility function who faces knightian uncertainty over a FOSD-ordered family of posteriors may restrict attention to increasing decision rules.

3.5 Informativeness

Consider the problem of choosing an action $x \in X$ based on a noisy signal $y \in Y$ of the state of the world $s \in S$, described in the previous section. Given two families P and P' of posteriors, P' is *more informative* than P for a decision maker with utility function $u(x, s)$ and signal distribution F if, for any $\psi \in D_F$, there exists $\phi \in D_F$ such that

$$\int_Y \int_S u(\phi(y), s)\mathrm{d}P'(s|y)\mathrm{d}F(y) \geq \int_Y \int_S u(\psi(y), s)\mathrm{d}P(s|y)\mathrm{d}F(y). \tag{14}$$

Intuitively, a more informative family of posteriors gives a decision maker higher ex-ante utility if she picks the right decision rule.

It is well-known that, if P is FOSD-ordered, then P' is more informative than P for any supermodular u and any F such that the follwing holds. The joint distribution over (y, s) induced by P' and F dominates the one induced by P and F in the *supermodular stochastic order* (written $P' \geq_F P$), i.e.

$$\int_Y \int_S v(y, s)\mathrm{d}P'(s|y)\mathrm{d}F(y) \geq \int_Y \int_S v(y, s)\mathrm{d}P(s|y)\mathrm{d}F(y) \tag{15}$$

41. Such preferences may arise in the presence of a status quo. For instance, if ψ is the status quo and ϕ is a potential alternative, ϕ is preferred to ψ if and only if it has higher value in all conceivable scenarios.

for all supermodular functions v.[42]

However, given ψ, the particular ϕ for which (14) holds is generally sensitive to the choice of u, P and P'. The next result states that rearranging a rule ψ with respect to F into an increasing rule ϕ raises the ex-ante utility of *all* decision makers with supermodular utility functions who switch from *any* posterior family P to *any* other family P' that satifies the above informativeness criterion.

Proposition 8. *Let $\psi \in D_F$ and let ϕ be the increasing rearrangement of ψ with respect to F. Then (14) holds for any supermodular u and any posterior families P' and P such that P is FOSD-ordered and $P' \geq_F P$.*

Proof. Fix $\psi \in D_F$, P and u. Since P is FOSD-ordered then, Proposition 7 implies that (12) holds. Moreover, since ϕ is increasing and u is supermodular, then the function $v(y,s) = u(\phi(y),s)$ is supermodular. Since $P' \geq_F P$ then (15) holds. Result follows by combining the inequalities. □

Proposition 8 improves on existing results in statistical decision theory. A *experiment* $G = \{G(\cdot|s)\}_{s\in S}$ is a collection of distributions $G(\cdot|s)$ of the signal y conditional on the value of the state s.[43] An experiment H is *more informative* than G for a decision maker with utility u if for any rule ψ there exists another rule ϕ such that $\int_Y u(\phi(y),y)dH(y|s) \geq \int_Y u(\psi(y),y)dG(y|s)$ *for all* $s \in S$.[44] In contrast, Proposition 8 states that, given F, for any ψ, there exists ϕ such that (14) holds for any P' such that $P' \geq_F P$. Results about informativeness of experiments cannot be used to derive Proposition 8 since, given P_1, P_1', P_2, P_2' and F such that $P_i' \geq_F P_i$ for $i = 1,2$, there might not exist experiments G and H such that H is more informative than G, and priors Λ_1, Λ_2 over S, such that G and Λ_i give rise to P_i and F and H and Λ_i give rise to P_i' and F for $i = 1,2$.

42. The supermodular stochastic order admits a useful characterisation. See Epstein and Tanny (1980), Meyer and Strulovici (2012) and Meyer and Strulovici (2015).

43. Note that there exists a one-to-one mapping between pairs (Λ, G), where Λ is a prior distribution over S, and pairs (P,F), which preserves the induced joint distribution over $Y \times S$.

44. Important informativeness criteria are due to Blackwell (1953) and Lehmann (1988). Blackwell's criterion captures informativeness when no substantial restrictions are imposed on experiments and utility functions, whereas Lehmann's criterion applies to statistical experiments that are MLR-ordered and quasiconcave families of utility functions with increasing peaks. Quah and Strulovici (2009) proved that Lehman's criterion applies to utilities satisfying the (weaker) interval dominance order property.

Example 1 – continued: Recall that a seller offers an insurance $\psi \in D_F$ inducing a net transfer $\psi(y_i)$ to agent i, where y_i is the realisation of a noisy signal of the shock s_i. The insurer minimises $\int_Y \psi(y)\mathrm{d}F(y)$ subject to the constraint that all agents are willing to buy the insurance. For agent i, this holds whenever $\int_Y \int_S u_i(\psi(y_i), s_i)\mathrm{d}P_i(s_i|y_i)\mathrm{d}F(y_i) \geq \bar{v}_i$ where u_i is supermodular.

Assume that $P_i = P$ for all i, where P is some FOSD-ordered family. This means that the agents are exposed to the same shock. However, risk attitutes (described by the curvatures of u_i's) may still differ across agents. Assume that P and F are obtained by the insurer from some experiment G and a prior Λ over S. Suppose that the insurer gains access to an experiment H that is more accurate than G in Lehmann's sense.[45] Without loss of generality, we may assume that H and Λ give rise to the same marginal distribution F for some posterior family P'.[46] Moreover, one can show that $P' \geq_F P$.[47] Then, Proposition 8 implies that, for any ψ, the increasing rearrangement of ψ with respect to F solves $\int_Y \int_S u_i(\phi(y), s)\mathrm{d}P'(s|y)\mathrm{d}F(y) \geq \int_Y \int_S u_i(\psi(y), s)\mathrm{d}P(s|y)\mathrm{d}F(y)$ for all i. Thus, agents who buy ψ when the insurer has access to G will also buy ϕ when the insurer has access to H. Moreover, since ϕ is a rearrangement of ψ, $\int_Y \phi(y)\mathrm{d}F(y) = \int_Y \psi(y)\mathrm{d}F(y)$. It follows that, if the insurer has access to H, she is able to sell an insurance to all agents inducing (weakly) lower expected transfers.[48]

Example 2 – continued: Recall that an ambiguity-averse firm commits to a decision rule ψ before observing a signal y of the state s. The firm knows that y has marginal distribution F, and that the family P of posteriors of s given y belongs to some set $\mathcal{P}$ of FOSD-ordered families. When the firm optimlises, its

45. The experiment H is more accurate than G in Lehmann's sense whenever there exists a map $T : Y \times S \to Y$ such that $T(y, \cdot)$ is increasing for all $y \in Y$, and $H(T(y, s)|s) = G(y|s)$ for all $y \in Y$ and $s \in S$.

46. Note that H and Λ give rise to some posterior family P' and some marginal F' which need not equal F. However, if F' is continuous, then the set of distributions over (x, s) that the insurer is able to induce is the same whether she has access to F' or to F.

47. See Milgrom (1981) for a proof.

48. This example is close to Example 6 in Quah and Strulovici (2009), in which a fund manager allocates the wealth of a group of heterogeneous investors between a safe and a risky asset, based on a signal of its return. They show that access to a signal that is more accurate in Lehamnn's sense allows the manager to increase the utility of all investors, irrespective of their increasing utility functions and priors. Their reasoning cannot be used to derive our result because it requires the utility functions to be increasing transformations of each other for any given s. In our setting, this is not the case because the insurer's utility is decreasing in x whereas the agents' utilities are increasing in x.

ex-ante utility $V(\mathcal{P})$ is given by (13).

Given two sets $\mathcal{P}$ and $\mathcal{P}'$, when is it the case that $V(\mathcal{P}') \geq V(\mathcal{P})$? We claim that this holds if, for any $P' \in \mathcal{P}'$ there exists $P \in \mathcal{P}$ such that $P' \geq_F P$. If this is so, then Proposition 8 implies that, for any ψ, there exists an increasing ϕ such that, for any $P' \in \mathcal{P}'$, (14) holds for some $P \in \mathcal{P}$. Since $P' \in \mathcal{P}'$ is arbitrary, then

$$\inf_{P' \in \mathcal{P}'} \int_Y \int_S u(\phi(y), y) \mathrm{d}P'(s|y) \mathrm{d}F(y) \geq \inf_{P \in \mathcal{P}} \int_Y \int_S u(\psi(y), y) \mathrm{d}P(s|y) \mathrm{d}F(y).$$

Since ψ is arbitrary, then $V(\mathcal{P}') \geq V(\mathcal{P})$.

4 Conclusion

In this paper, we showed that it is possible to rearrange any decision rule into an increasing rule in a way that does not alter the induced distribution of actions. Moreover, doing so increases the expected payoff induced by any supermodular utility function. In particular, for the purposes of the simultaneous maximisation of supermodular expected payoffs, the decision maker may restrict attention to increasing decision rules.

Appendix

Note that any $\psi \in D_F$ induces a distribution over $X \times Y$ given by

$$J_{\psi F}(x, y) = \int_Y \mathbb{I}\{y' \leq y\} \psi\left(x|y'\right) \mathrm{d}F\left(y'\right). \tag{16}$$

The following facts will be used in the proof of Lemma 1.

Lemma 2. *Let $\psi \in D_F$ be increasing and comonotone. Then $J_{\psi F}$ is comonotone.*

Proof. Pick $(x, y), (\hat{x}, \hat{y}) \in X \times Y$ in the support of $J_{\psi F}$. If $x = \hat{x}$, then either $(x, y) \leq (\hat{x}, \hat{y})$ or $(x, y) \geq (\hat{x}, \hat{y})$ and result holds. Therefore, suppose without loss of generality that $\hat{x}_1 > x_1$. Then it suffices to show that $(\hat{x}, \hat{y}) \geq (x, y)$. Since (x, y) is in the support of $J_{\psi F}$, $J_{\psi F}$ assigns positive probability to all neighbourhoods of (x, y). Hence, there exists a sequence $(x^m, y^m)_m \subset X \times Y$ converging to (x, y)

such that x^m is in the support of $\psi(\cdot|y^m)$ for all $m \geq 0$. Similarly, there exists a sequence $(\hat{x}^m, \hat{y}^m)_m \subset X \times Y$ converging to $(\hat{x}, \hat{y})$ such that $\hat{x}^m$ is in the support of $\psi(\cdot|\hat{y}^m)$ for all $m \geq 0$. If, for some $i \neq 1$, $\hat{x}_i < x_i$, then there would exist $m \geq 0$ such that $\hat{x}_i^m < x_i^m$ and $\hat{x}_1^m > x_1^m$. But then $\psi(\cdot|y^m)$ would not be comonotone. Hence $\hat{x} \geq x$. If $\hat{y} < y$, there would exist $m \geq 0$ such that $\hat{y}^m < y^m$ and $\hat{x}_1^m > x_1^m$. Then $\psi(x'|\hat{y}^m) < 1$ and $\psi(x'|y^m) > 0$, where $x_1' = (\hat{x}_1^m + \hat{x}^m)/2$ and $x_i' = \max X_i$ for $i > 1$. This contradicts the fact that ψ is increasing. Hence $\hat{y} \geq y$. Result follows. □

Proof of Lemma 1. I prove that ϕ is unique. Suppose that ϕ' is another increasing and comonotone rearrangement of ψ with respect to F. Then it suffices to show that $\phi(\cdot|y) = \phi'(\cdot|y)$ for F-almost all y's. Since ϕ and ϕ' are increasing and comonotone, then Lemma 2 implies that $J_{\phi F}$ and $J_{\phi' F}$ are comonotone. Moreover, since ϕ and ϕ' are rearrangements of ψ, $J_{\phi F}$ and $J_{\phi' F}$ have identical one-dimensional marginal distributions. Then, since $J_{\phi F}$ and $J_{\phi' F}$ are comonotone, $J_{\phi F} = J_{\phi' F}$. Suppose that there exists a set $Y_0 \subset Y$ of positive F-measure such that $\phi(\cdot|y) \neq \phi'(\cdot|y)$ for all $y \in Y_0$ and seek a contradiction. If this is so then, without loss of generality, for any $y \in Y_0$ there exists $x \in X$ such that $\phi(x|y) > \phi'(x|y)$. Then there exists $Y_1 \subseteq Y_0$ and $\epsilon > 0$ such that, for all $y \in Y_1$ there exists $x \in X$ such that $\phi(x|y) - \phi'(x|y) > \epsilon$. Let $g : Y_1 \to X$ be given by $g(y) = \inf\{x \in X : \phi(x|y) - \phi'(x|y) > \epsilon\}$. Since $X = X_1 \times \cdots \times X_n$ where $X_i \subset \mathbb{R}$ is compact for all i, g is well-defined. Since $\phi(\cdot|y)$ and $\phi'(\cdot|y)$ are right-continuous, $\phi(g(y)|y) - \phi'(g(y)|y) \geq \epsilon$ for all $y \in Y_1$. Moreover, it is clear that g is measurable. It follows that $\int_{Y_1} \phi(g(y)|y)\mathrm{F}(y) - \int_{Y_1} \phi'(g(y)|y)\mathrm{F}(y) \geq \epsilon \int_{Y_1} \mathrm{d}F(y) > 0$. That is, $J_{\phi F}$ puts larger mass on $\{(x, y) \in X \times Y_1 : x \leq g(y)\}$ than $J_{\phi' F}$. This contradicts the fact that $J_{\phi F} = J_{\phi' F}$. □

Theorem 1 follows from a version of Lorentz's inequality proved by Burchard and Hajaiej (2006). Proposition 9 below is a corollary of their result.

Proposition 9 (Corollary of Burchard and Hajaiej 2006)**.** *Given $\varphi : (0, 1] \to \mathbb{R}^m$, let $\varphi' : (0, 1] \to \mathbb{R}^m$ be such that*

$$\varphi_i'(\omega) = \inf\{z \in \mathbb{R} : \mu(\{\omega' \in (0, 1] : \varphi_i(\omega') \leq z\}) \geq \omega\} \tag{17}$$

for all $i \leq m$. *Then*

$$\int_0^1 v\left[\varphi'(\omega)\right] d\omega \geq \int_0^1 v[\varphi(\omega)] d\omega \tag{18}$$

for any supermodular function $v : \mathbb{R}^m \to \mathbb{R}$ *such that*

$$\int_0^1 \sup_{z \in \mathbb{R}^{m-1}} |v(z, \varphi_m(\omega))| d\omega < \infty. \tag{19}$$

Moreover, unless

$$[\varphi_i(\omega) - \varphi_i(\omega')][\varphi_j(\omega) - \varphi_j(\omega')] \geq 0 \tag{20}$$

for all $i, j \leq m$ *and almost all* $\omega, \omega' \in (0, 1]$, (18) *is strict for any strictly supermodular* v *such that* (19) *holds.*[49]

The *generalised inverse* of a one-dimensional distribution G is the function $G^- : (0, 1] \to \mathbb{R}$ given by

$$G^-(\omega) = \inf\{z \in \mathbb{R} : G(z) \geq \omega\}.$$[50]

The following fact will be used in the proof of Theorem 1.

Lemma 3. *Let* G *be a comontone distribution and let* $G_1^-, \ldots, G_m^-$ *be the generalised inverses of its one-dimensional marginals. Let* $\varphi : (0, 1] \to \mathbb{R}^m$ *be given by* $\varphi(\omega) = [G_1^-(\omega), \ldots, G_m^-(\omega)]$. *Then* φ *is a random vector with distribution* G.

Proof. Note that $\varphi_0 : (0, 1] \to \mathbb{R}$ given by $\varphi_0(\omega) = \omega$ is a random vector with uniform distribution over $[0, 1]$. Then result follows from the implication (3) $\Rightarrow$ (1) of Theorem 2 of Dhaene et al. (2002). □

Proof of Theorem 1. To prove the first part, fix $\psi \in D_F$ and let ϕ be the increasing and comonotone rearrangement of ψ with respect to F. Then it suffices to show that (3) holds for any $u \in \mathcal{U}_{\text{SPM}}$ such that $\int_Y \sup_{x \in X} |u(x, y)| \mathrm{d}F(y) < \infty$. Fix

49. Proposition 9 is a special case of a corollary of Theorem 1 of Burchard and Hajaiej (2006), as expressed in equation (2.2). The authors use a different notation: our φ (resp. φ', v) corresponds to their u (resp. $u^\#$, F). Moreover, they require that φ is positive as they impose an integrability condition on v (expressed in equation 2.1) that is weaker than (19). Given our stronger assumption, the requirement that φ is positive is void as one may replace φ_i with e^{φ_i} and v with $v[\log(\cdot), \ldots, \log(\cdot)]$. Note also that, since the domain of φ is restricted to $(0, 1]$, φ trivially vanishes at infinity. Finally, the second part of Proposition 9 is not stated explicitly in Burchard and Hajaiej (2006), but it follows easily from the second part of their Theorem 1.

50. If G is invertible, then its inverse equals G^-.

u and let $\varphi : (0,1] \to X \times Y$ be a random vector with distribution $J_{\psi F}$, defined in (16). Then (19) holds. Hence Proposition 9 implies that (18) holds, where $\varphi' : (0,1] \to \mathbb{R}^{n+1}$ is defined by (17). Note that, since φ has distribution $J_{\psi F}$ then $\int_Y \int_X u(x,y)\mathrm{d}\psi(x|y)\mathrm{d}F(y) = \int_0^1 u[\varphi(\omega)]\mathrm{d}\omega$. Then, it suffices to show φ' is a random vector with distribution $J_{\phi F}$. Since ϕ is increasing and comonotone, then Lemma 2 implies that $J_{\phi F}$ is comonotone. Moreover, for any $z \in \mathbb{R}$, $\mu(\{\omega' \in (0,1] : \varphi_i(\omega') \leq z\}) = M_{\psi_i F}(z)$ for all $i \leq n$, and $\mu(\{\omega' \in (0,1] : \varphi_{n+1}(\omega') \leq z\}) = F(z)$. Then φ_i' is the generalised inverse of $M_{\psi_i F}$ for all $i \leq n$, and φ_{n+1} is the generalised inverse of F. Finally, since ϕ is a rearrangement of ψ, the one-dimensional marginal distributions of $J_{\psi F}$ and $J_{\phi F}$ coincide. Hence Lemma 3 implies that φ is a random vector with distribution $J_{\phi F}$. Result follows.

To prove the second part, note that, from Lemma 1, the increasing and comonotone rearrangement of ψ with respect to F is unique. Therefore, it suffices to show that, if ψ is $\mathcal{U}_{\mathrm{SPM}}$-dominated by some rule $\phi \in D_F$, then ϕ is a rearrangement of ψ with respect to F. Fix $\phi, \psi \in \hat{D}_F$ such that ϕ is not a rearrangement of ψ with respect to F. Then it remains to show that ϕ does not $\mathcal{U}_{\mathrm{SPM}}$-dominate ψ, i.e. that (3) fails for some supermodular payoff u. Since ϕ is not a rearrangement of ψ, $M_{\phi_i F}(x_i') \neq M_{\psi_i F}(x_i')$ for some $i \leq n$ and $x_i' \in X_i$. If $M_{\phi_i F}(x_i') > M_{\psi_i F}(x_i')$, then pick $u(x,y) = \mathbb{I}\{x_i > x_i'\}$. If $M_{\phi_i F}(x_i') < M_{\psi_i F}(x_i')$, then pick $u(x,y) = \mathbb{I}\{x_i \leq x_i'\}$. Note that u is supermodular in either case. Hence ϕ does not $\mathcal{U}_{\mathrm{SPM}}$-dominate ψ. Result follows.

To prove the last part, suppose that ψ is not strictly $\mathcal{U}_{\mathrm{SSPM}}$-dominated by its increasing and comonotone rearrangement with repsect to F, ϕ. It suffices to show that ψ is increasing and comonotone. Note that there exists a strictly supermodular payoff function u such that (3) is not strict. The argument to prove the first part of the theorem implies that (18) is not strict. Then, Proposition 9 implies that (20) holds for all $i, j \leq n+1$ and almost all $\omega, \omega' \in (0,1]$. Since φ has distribution $J_{\psi F}$, it follows that $J_{\psi F}$ is comonotone.[51] Then, the argument used to prove uniqueness in Lemma 1 implies that $\phi(\cdot|y) = \psi(\cdot|y)$ for F-almost all y's. Hence ψ is increasing and comonotone. □

51. This follows from the definition of comonotonicty, see Section 2.

Proof of Proposition 1. Since X has more than one element, X_i has more than one element for some $i \leq n$. Without loss of generality, assume that $\{0,1\} \in X_i$. Given $y_0 < y_1$, let $\psi \in D_F$ be such that $\psi_i(y) = 1 - \mathbb{I}\{y_0 < y \leq y_1\}$. Since there exists an increasing rearrangement ϕ of ψ with respect to F, $F(y_1) - F(y_0) = M_{\psi_i F}(0) = M_{\phi_i F}(0) \in \overline{F(\mathbb{R})}$, where the inclusion follows from the fact that ϕ is increasing. Then, it suffices to show that, unless F is a continuous or a discrete uniform distribution, $F(y_1) - F(y_0) \notin \overline{F(\mathbb{R})}$ for some $y_0 < y_1$. Suppose that F does not have an atom at the bottom of its support. Formally, suppose that for all $\epsilon > 0$, there exists $y_0 \in \mathbb{R}$ such that $F(y_0) \in (0, \epsilon)$. Since F is not a continuous distribution, it has an atom $y_1 \in \mathbb{R}$. Then there exists $y_0 \in \mathbb{R}$ such that $F(y_0) \in (0, p)$ where p is the probability mass at y_1. Hence $y_0 < y_1$ and $F(y_1) - F(y_0) \notin \overline{F(\mathbb{R})}$. Suppose instead that F has an atom y_0 at the bottom of its support. Since F is not a discrete uniform distribution, there exists $y_1 \in \mathbb{R}$ such that $F(y_1) \notin \{mF(y_0) : m \in \mathbb{N}_0, m \leq F(y_0)^{-1}\}$. Then $y_1 > y_0$. Moreover, y_1 can be chosen such that $F(y_1) - F(y_0) \notin \overline{F(\mathbb{R})}$. □

Proof of Proposition 2. Suppose that $u \in \mathcal{U}$ is not supermodular. It suffices to find a distribution F and $\psi \in D_F$ such that (3) fails if $\phi \in D_F$ is an increasing rearrangement of ψ with respect to F. Since u is not supermodular, there exists $(x^1, y^1), (x^2, y^2) \in X \times Y$ such that $y^2 \geq y^1$ and $u(x^1 \wedge x^2, y^1) + u(x^1 \vee x^2, y^2) < u(x^1, y^1) + u(x^2, y^2)$. Let F be the uniform distribution over $\{y_1, y_2\}$ and let $\psi \in \hat{D}_F$ be such that $\psi(y^j) = x^j$ for $j = 1, 2$. Then, ψ admits a unique increasing rearrangement with respect to F, given by $\phi(y^1) = x^1 \wedge x^2$ and $\phi(y^2) = x^1 \vee x^2$. So

$$\begin{aligned}
&\int_Y u(\phi(y), y) \mathrm{d}F(y) - \int_Y u(\psi(y), y) \mathrm{d}F(y) \\
&= u(x^1 \wedge x^2, y^1) - u(x^1, y^1) + u(x^1 \vee x^2, y^2) - u(x^2, y^2) < 0.
\end{aligned}$$

Therefore (3) fails. □

References

Arieli, I., and Y. Babichenko. 2016. "Private bayesian persuasion." Working paper, available at: https://ssrn.com/abstract=2721307.

Athey, S., and J. Levin. 1998. "The value of information in monotone decision problems." MIT Working Paper.

Bewley, T. 2002. "Knightian decision theory. Part I." *Decisions in economics and finance* 25 (2): 79–110.

Blackwell, D. 1953. "Equivalent comparisons of experiments." *The annals of mathematical statistics* 24 (2): 265–272.

Bloch, F., and N. Quérou. 2013. "Pricing in social networks." *Games and economic behavior* 80:243–261.

Burchard, A., and H. Hajaiej. 2006. "Rearrangement inequalities for functionals with monotone integrands." *Journal of Functional Analysis* 233 (2): 561–582.

Candogan, O., K. Bimpikis, and A. Ozdaglar. 2012. "Optimal pricing in networks with externalities." *Operations Research* 60 (4): 883–905.

Cerreia-Vioglio, S., D. Dillenberger, and P. Ortoleva. 2015. "Cautious expected utility and the certainty effect." *Econometrica* 83 (2): 693–728.

Cerreia-Vioglio, Simone. 2009. "Maxmin expected utility on a subjective state space: Convex preferences under risk."

Chernozhukov, V., I. Fernandez-Val, and A. Galichon. 2009. "Improving point and interval estimators of monotone functions by rearrangement." *Biometrika* 96 (3): 559–575.

Dhaene, Jan, Michel Denuit, Marc J Goovaerts, Rob Kaas, and David Vyncke. 2002. "The concept of comonotonicity in actuarial science and finance: theory." *Insurance: Mathematics and Economics* 31 (1): 3–33.

Dziewulski, P., and J. Quah. 2016. "Supermodular Correspondences." University of Oxford Working Paper, no. 795.

Epstein, L., and M. Le Breton. 1993. "Dynamically consistent beliefs must be Bayesian." *Journal of Economic theory* 61 (1): 1–22.

Epstein, L., and M. Schneider. 2003. "Recursive multiple-priors." *Journal of Economic Theory* 113 (1): 1–31.

Epstein, L., and S. Tanny. 1980. "Increasing generalized correlation: a definition and some economic consequences." *Canadian Journal of Economics:* 16–34.

Fainmesser, I., and A. Galeotti. 2015. "Pricing network effects." *The Review of Economic Studies.*

Farrell, J., and G. Saloner. 1985. "Standardization, compatibility, and innovation." *The RAND Journal of Economics:* 70–83.

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

———. 1993. "Updating Ambiguous Beliefs." *Journal of Economic Theory* 59 (1): 33–49.

Hanany, E., and P. Klibanoff. 2007. "Updating preferences with multiple priors." *Theoretical Economics* 2 (3): 261–298.

Karlin, S., and H. Rubin. 1956. "The theory of decision procedures for distributions with monotone likelihood ratio." *The Annals of Mathematical Statistics:* 272–299.

Katz, M. L., and C. Shapiro. 1985. "Network externalities, competition, and compatibility." *The American economic review* 75 (3): 424–440.

Landsberger, M., and I. Meilijson. 1994. "Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion." *Annals of Operations Research* 52 (2): 97–106.

Lehmann, E. L. 1988. "Comparing Location Experiments." *The Annals of Statistics:* 521–533.

Lorentz, G. 1953. "An inequality for rearrangements." *The American Mathematical Monthly* 60 (3): 176–179.

Meyer, M. 2018. "Information Design: Insights from Orderings of Dependence and Heterogeneity."

Meyer, M., and B. Strulovici. 2015. *Beyond correlation: Measuring interdependence through complementarities.* Technical report. Working Paper.

———. 2012. "Increasing interdependence of multivariate distributions." *Journal of Economic Theory* 147 (4): 1460–1489.

Milgrom, P. 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics:* 380–391.

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

Mirrlees, J. A. 1971. "An exploration in the theory of optimum income taxation." *The review of economic studies* 38 (2): 175–208.

Quah, J. K.-H., and B. Strulovici. 2009. "Comparative statics, informativeness, and the interval dominance order." *Econometrica* 77 (6): 1949–1992.

Ramsey, F. 1927. "A Contribution to the Theory of Taxation." *The Economic Journal* 37 (145): 47–61.

Stiglitz, J. 1982. "Utilitarianism and horizontal equity: The case for random taxation." *Journal of Public Economics* 18 (1): 1–33.

Sundararajan, A. 2007. "Local network effects and complex network structure." *The BE Journal of Theoretical Economics* 7 (1).

Chapter 2

Incentives for Collective Innovation

Symmetric Equilibrium with Full Disclosure

Abstract

Identical agents exert hidden effort to produce random increments in the stock of a public good, and receive a stock- and effort-contingent flow of payoffs. I characterise the social-welfare benchmark, as well as the unique symmetric Markov equilibrium of the game, for a broad class of payoff functions. In the welfare benchmark, any increment in the stock is beneficial. However, if the opportunity cost of effort increases sufficiently fast with the stock, equilibrium continuation payoffs may drop after an innovation. In the context of an R&D partnership in which improvements in a shared technology raise the incentive to shift resources towards private activities, moderate improvements may be counter-productive.

1 Introduction

I analyse a family of dynamic games in which N identical agents exert hidden effort to induce increments in the stock of a public good. The games have two novel features. First, not only is the arrival of increments stochastic (their frequency being increasing with effort) but so is their size. Second, agents' flow payoffs are a general function of their effort and of the current stock, and the marginal cost of effort is (weakly) increasing in the stock. This class of games captures instances of collective innovation in which progress decreases the incentive to innovate further.[1]

1. My motivation has much in common with that of the literature on dynamic games of public-good provision, from Bliss and Nalebuff (1984) to Battaglini, Nunnari, and Palfrey (2014) (see Sections 1.1 of

As an example, consider a group of firms improving a common technology and sharing their innovations.[2] Knowledge-sharing among innovators played a major role in the biotechnology industry (Powell, Koput, and Smith-Doerr 1996). Historically, this phenomenon contributed to the development of the blast furnace (Allen 1983) and steam-operated mining devices (Nuvolari 2004) in the UK, and the personal computer in the US (Meyer 2003). These instances of 'collective invention' are characterised by the accumulation of many technological improvements, obtained by different actors over a period of time. In this setting, the stock of the public good corresponds to the current technology level, and an increment in the stock is an innovation. Firms carry out R&D to obtain innovations of random magnitudes and frequency, as well as other activities (e.g. production, marketing, or product development) inducing a private flow of profits.[3] Profits are increasing in the technology level and, if firms are resource-constrained, then the marginal opportunity cost of R&D is increasing, too.[4]

This chapter has two main results. First, I characterise the social-welfare benchmark in a general enviroment. I show that, in order to maximise welfare, agents should exert less effort as the stock grows. I also derive a necessary and sufficient condition for effort to vanish in the limit as the stock diverges.

Second, I construct the unique symmetric Markov-perfect equilibrium with the stock as the state variable. Several properties of the symmetric equilibrium are common in dynamic public-good games. Namely, effort decreases as the stock grows and it is inefficiently low due to the 'inter-temporal free-riding' effect (Marx and Matthews 2000). Moreover, effort ceases when it would cease in the single-agent case (i.e. the encouragement effect found in Bolton and Harris (1999) does not arise).

this chapter and of Chapter 3 for a review).

2. Either an R&D partnership or an informal disclosure agreement would fit the model.

3. Assume that firms' cannot observe each other's R&D efforts, and are thus subject to moral hazard. Geroski (1995) notes that 'co-operative R&D ventures may be unable to overcome moral hazard problems that lead participants to invest less than promised in the joint venture, divert the energies of people nominally assigned to [it, and] assign less talented researchers to [it...]'

4. Powell and Giannella (2010) note that progress may hinder R&D incentives: 'Once knowledge accumulates to the stage that tangible outcomes are possible, private interests may take hold and commercialise particular streams of technology that emerge from collective invention [...] as technological uncertainty recedes, firms develop private R&D and focus on their own specific applications. Reliance on collective invention [...] wanes.'

The main novelty of the equilibrium is that, under general conditions, continuation payoffs may fall after an innovation (i.e. an increment in the stock). This is because effort decreases as the stock grows, delaying future innovations, and this may offset the benefit of a moderately larger stock. In contrast, a sufficiently large innovation is always beneficial, even though it leads to a sharp drop in effort.

A sufficient condition for payoffs to be initially decreasing in the stock is that the marginal cost of effort is sufficiently low when the stock is sufficiently small, and there are many agents. This is because, for any given stock, equilibrium individual effort is low if the number of agents is large. Then, if the stock is small, individual equilibrium marginal cost is negligible. Since agents optimise, the marginal benefit must be negligible, too. Hence, payoffs are approximately the *average* of payoffs after the first (randomly-sized) innovation. Since payoffs are not constant, it follows that they cannot be increasing in the stock, otherwise payoffs before the innovation would be strictly lower than this average. In contrast, in the social-welfare benchmark, payoffs are always increasing in the stock.[5]

In the context of innovation-sharing among firms, if the returns from private activities are negligible when the shared technology is very inefficient, and there are many firms, then small innovations may be harmful.

This class of dynamic-stochastic games is mathematically novel as the state (i.e, the stock) makes randomly-sized jumps, so that agents' value functions solve integral equations. To characterise the social-welfare benchmark and the symmetric equilibrium, I extend considerably the arguments of Bolton and Harris (1999).

1.1 Literature review

This paper contributes to the literature of dynamic public-good games by showing that, when production occurs in lumps of random size, equilibrium continuation payoffs may fall after an increment in the stock. This is not true in the widely studied case of deterministic production (e.g. Admati and Perry 1991, Gradstein (1992), and Fershtman and Nitzan 1991), nor if production is stochastic but con-

5. In particular, this analysis does not apply since *aggregate* marginal cost is not negligible.

tinuous, such as in Wang and Ewald (2010).[6]

The game is also related to the strategic experimentation literature, and in particular to Bolton and Harris (1999), Keller, Rady, and Cripps (2005), Keller and Rady (2010), Bonatti and Hörner (2011), and Bimpikis, Drakopoulos, and Ehsani (2018). However, in contrast to this model, the production of 'good news' in these games (which may be thought of as an increment in the stock) *increases* the incentive to experiment (i.e. to exert effort).

Finally, the game is related to the large theoretical literature on innovation. Katz (1986) studies to what extent co-operative R&D may spur innovation among firms competing in the product market, based on the degree of competition, the amount of R&D spillovers, and the type of co-operative research undertaken. Dasgupta (1988) argues that the first-mover disadvantage may give rise to a waiting game among innovators in the presence of large knowledge spillovers and high R&D costs. Taylor (1995) shows that, when designing contests for innovation, entry barriers may be beneficial. Cowan and Jonard (2003) model information spillovers with a network, and analyse how its architecture influences the patterns and the rate of innovation.

The rest of the chapter is organised as follows. Section 2 contains a description of the model. In Section 3, I present the social-welfare benchmark. In Section 4, I characterise the unique symmetric Markov-perfect equilibrium, and analyse the impact of innovations on continuation payoffs. The conclusion is in Section 5.

2 Model and Assumptions

In this section, I describe the general model and introduce a specific example which will be useful to illustrate the results. The section ends with a discussion of the assumptions.

Framework. Time is continuous and indexed by $t \in [0, \infty)$. Each of N agents, indexed by $i \leq N$, exerts costly effort to produce a public good. Let $x_t \geq 0$ denote

6. In fact, the effect disappears even in this model if the size of lumps is deterministic, or in the limit as production becomes continuous.

the stock of the good at time $t \geq 0$. In any interval of time $[t, t+\mathrm{d}t)$, Agent i exerts effort $a_t^i \in [0,1]$ and receives payoff $u(a_t^i, x_t)$, discounted at rate $r > 0$. Assume that u is twice continuously differentiable, increasing and concave in x.[7] That is, keeping the effort fixed, payoffs increase as the stock grows, but at a decreasing rate. Note that payoffs u may be expressed as $u(a,x) = b(x) - c(a,x)$, where $b(x) = u(0,x)$ are gross benefits, and $c(a,x) = u(0,x) - u(a,x)$ is the opportunity cost of effort. Suppose that $c(a,x)$ is increasing and convex in a, and strictly increasing in a if $x > 0$. Suppose also that $c_1(a,x)$ and $c_{11}(a,x)$ are increasing in x; that is, as the stock grows, the marginal cost of effort rises, and the cost becomes more convex.[8] Finally, assume for simplicity that, for any $x \geq 0$, $c_{11}(a,x)$, $c_{112}(a,x)$ are bounded as a tends to 1, and that $b'(x)$ is bounded as x shrinks to 0.[9]

The stock x_t takes some initial value $x_0 \geq 0$ and, for $t > 0$, is determined as follows. In any interval of time $[t, t+\mathrm{d}t)$, for any $i \leq N$, Agent i produces an increment in x_t with probability $a_t^i \mathrm{d}t$. Each increment has random size s, drawn from a CDF F with $F(0) = 0$, and mean $\mu_F < \infty$. The production and the size of increments are independent across agents, time, and from one another. For most results, I assume that $\lim_{x\to\infty} \mu_F b'(x) - r c_1(a,x) < 0$ for all $a > 0$.

Rigorous definitions of strategies, histories, and payoffs are in Section **??**. In the main interpretation of the model, the stock x_t denotes the level of a technology, and increments in x_t are *innovations*. These terms are used throughout the discussion. Moreover, the following special case is used to illustrate the results.

Example. Suppose that agents are firms sharing a production technology, and that they may carry-out R&D to make it more efficient. The stock x_t denotes the current technology level, and increments in the stock are innovations. In each period, firms may either carry out R&D ($a_t^i = 1$), or focus on production ($a_t^i = 0$). Production induces a flow of profits $b(x_t) = x_t$, and R&D has an opportunity cost

7. Throughout the discussion, notions such as 'increasing' and 'concave' are meant in the weak sense.

8. The conditions on c are equivalent to the following conditions on u: 'u is decreasing and concave in a, and strictly decreasing in a if $x > 0$; u_1 and u_{11} are decreasing in x'.

9. This rules out payoffs $b(x) = \sqrt{x}$ and $b(x) = \log(x)$, and cost functions such as $c(a,x) = 1 - \sqrt{1-a}$. Relaxing these assumptions does not alter the results but requires more technical proofs.

$c(1, x_t) = \alpha x_t$, for some $\alpha \in (0, 1]$. Hence

$$u(a, x) = \begin{cases} x & \text{if } a = 0 \\ x(1 - \alpha) & \text{if } a = 1. \end{cases}$$

In particular, production becomes more efficient as the technology level increases. If firms carry out R&D with probability $a \in (0, 1)$, their expected payoff is $u(a, x) = x(1 - \alpha a)$.

R&D generates innovations after exponentially-distributed random times. The distribution of the size of innovations is $F(x) = 1 - e^{-\lambda x}$ for some $\lambda > 0$.[10]

Discussion of the assumptions. The literature on dynamic public-good games assumes that $u(a, x) = b(x) - c(a)$, where b is concave with $\lim_{x\to\infty} b'(x) = 0$, and c is convex. This model is a generalisation, since the cost c is a function of the stock x.

The requirement that $c_1(a, x)$ and $c_{11}(a, x)$ increase in x is a substantial one.[11] If the assumption that $c_{11}(a, x)$ is increasing in x is relaxed, then the optimal effort schedule need not be decreasing (c.f. Section 3, footnote 14).

The assumption that $\lim_{x\to\infty} \mu_F b'(x) - r c_1(a, x) < 0$ for all $a > 0$ means that exerting any fixed amount of positive effort becomes unprofitable as the technology level diverges. This is a generalisation of the condition that $\lim_{x\to\infty} b'(x) = 0$. This requirement implies that optimal effort vanishes as the technology level diverges (c.f. Section 3, Theorem 1). If the assumption is dropped, the other features of the welfare benchmark are unchanged, but the symmetric equilibrium (c.f. Theorem 2) may no longer be unique.

If an agent exerts constant effort a, the resulting arrival of innovations is a Poisson process with rate a. Since the arrival of innovations is independent across players, the processes aggregate. Namely, if agents exert constant effort $a^1, \ldots, a^N$, the overall arrival of innovations is a Poisson process with rate $\sum_{i=1}^N a^i$. Note that, given $\beta, \gamma > 0$, modifying the probability $a_t^i \mathrm{d}t$ of obtaining an innovation within the

10. Closed-form solutions are considerably harder to obtain if this assumption is dropped.

11. Note that, since $c_1(a, x)$ is increasing in x, $c(a, x)$ is increasing in x, too. Namely, since $c(0, x) = 0$, $c_2(a, x) = c_2(a, x) - c_2(0, x) = \int_0^a c_{12}(s, x)\mathrm{d}s \geq 0$.

time interval $[t, t+\mathrm{d}t)$ to $\beta a_t^i \mathrm{d}t$, and resizing the domain of a_t^i to $[0, \gamma]$ for all $i \leq N$, is strategically equivalent to replacing r with $r/(\beta\gamma)$ and $u(a, x)$ with $u(a/\gamma, x)$.

Finally, the distribution of the random size s of increments is independent of the effort a of the agent inducing it. This is a substantial assumption. However, it is without loss of generality if effort is binary; that is, if picking $a \in (0, 1)$ is viewed as a randomising between $a = 0$ and $a = 1$.

3 Social-Welfare Benchmark

In this section, I describe how non-strategic agents should behave in order to maximise aggregate expected discounted payoffs. The main features of this benchmark case are common in dynamic public-good games. In particular, any innovation is beneficial.

The model laid out in the previous section is a stochastic game with the stock x_t as the state variable. A *Markov* strategy is a strategy such that effort is pinned down by the current stock. Formally, it is a (measurable) function $\sigma^i : [0, \infty) \to [0, 1]$ such that $a_t^i = \sigma^i(x_t)$ for all $t \geq 0$.[12] The *initial payoffs* (resp. *continuation payoffs*) of Agent i (at time $t > 0$) are the expected discounted payoffs of Agent i at the beginning of the game (resp. at time t). Note that a profile $\sigma = (\sigma^i)_{i=1}^N$ of Markov strategies pins down the initial payoffs $v^i(x_0)$ of Agent i in the game with initial stock x_0 for any $x_0 \geq 0$. Moreover, when σ is played, the continuation payoffs of Agent i at time t equal $v^i(x_t)$. To derive an expression for $v^i(x_t)$, note that, for a small interval of time $\mathrm{d}t$,

$$\begin{aligned} v^i(x_t) &= r\left[b(x_t) - c\left(\sigma^i(x_t), x_t\right)\right]\mathrm{d}t + e^{-r\mathrm{d}t}\mathbb{E}\left[v^i(x_{t+\mathrm{d}t})\right] \\ &= r\left[b(x_t) - c\left(\sigma^i(x_t), x_t\right)\right]\mathrm{d}t + (1 - r\mathrm{d}t)\left[\textstyle\sum_{j=1}^N \sigma^j(x_t)\mathrm{d}t \int_0^\infty v^i(x_t + s)\mathrm{d}F(s)\right. \\ &\quad \left. + \left(1 - \textstyle\sum_{j=1}^N \sigma^j(x_t)\mathrm{d}t\right) v^i(x_t)\right] \end{aligned} \tag{1}$$

where $\sum_{j=1}^N \sigma^j(x_t)\mathrm{d}t$ is the probability that the stock increases within $[t, t + \mathrm{d}t)$, and $\int_0^\infty v^i(x_t + s)\mathrm{d}F(s)$ are the expected continuation payoffs of Agent i after an

12. By a slight abuse of notation, for symmetric profiles $(\sigma^i)_{i=1}^N$, σ refers both to the profile and to each strategy σ^i.

increment of random size s. Equation (1) follows by taking first-order approximations of $e^{-r\mathrm{d}t}$ and $\mathbb{E}\left[v^i(x_{t+\mathrm{d}t})\right]$ at $\mathrm{d}t = 0$. Neglecting terms in $\mathrm{d}t^2$, and dropping the subscript t, the above is equivalent to

$$rv^i(x) = r\left[b(x) - c\left(\sigma^i(x), x\right)\right] + \sum_{j=1}^{N} \sigma^j(x) \int_0^\infty v^i(x+s) - v^i(x)\mathrm{d}F(s). \quad (2)$$

Equation (2) states that the time-t continuation payoffs of Agent i are the sum of the current payoff flow $b(x) - c\left(\sigma^i(x), x\right)$ where $x = x_t$, and of the net expected future benefits. These are given by the rate $\sum_{j=1}^N \sigma^j(x)$ at which innovations occur, times their next expected social value; that is, the the difference between the continuation payoffs $v^i(x+s)$ after an innovation of size $s \geq 0$, and the current payoffs $v^i(x)$, weighted by the distribution F of s.

Equation (2) can be expressed as

$$v^i(x) = \frac{r\left[b(x) - c\left(\sigma^i(x), x\right)\right]}{r + \sum_{j=1}^N \sigma^j(x)} + \frac{\sum_{j=1}^N \sigma^j(x)}{r + \sum_{j=1}^N \sigma^j(x)} \int_0^\infty v^i(x+s)\mathrm{d}F(s). \quad (3)$$

In this case, continuation payoffs are the weighted average of the current payoff flow $b(x) - c\left(\sigma^i(x), x\right)$, and of expected continuation payoffs after an innovation. The weight on current payoffs increases with the time discount factor r, whereas the weight on payoffs after an innovation increases in the rate $\sum_{j=1}^N \sigma^j(x)$ at which innovations arrive. Both (2) and (3) will be used throughout the discussion.

Welfare is the average of agents' initial payoffs, and is given by

$$w(x) = \sum_{i=1}^{N} v^i(x).$$

Since this is a stochastic game, not only is welfare maximised by a profile of Markov strategies, but there exists a profile σ_* maximising welfare for all initial stocks x_0. Note that agents are identical, the cost of effort is convex, and the rate of arrival of innovations is linear in effort. Then, for any strategy profile σ, the symmetrised version of σ, i.e. the symmetric profile $\hat{\sigma}$ such that $\hat{\sigma}^i(x) = \frac{1}{N}\sum_{j=1}^N \sigma^j(x)$ for all $x \geq 0$, induces higher welfare. From (2), the Bellman equation for welfare

maximisation subject to effort being symmetric is

$$rw_*(x) = \max_{a\in[0,1]} r[b(x) - c(a,x)] + Na \int_0^\infty w_*(x+s) - w_*(x)\mathrm{d}F(s). \tag{4}$$

Therefore, if (4) admits a (well-behaved) solution w_*, then $w_*(x_0)$ is the maximal welfare achievable in the game with initial stock x_0, for all $x_0 \geq 0$. Moreover, the symmetric efficient effort schedules σ_* are the symmetric profiles such that, for all $x \geq 0$,

$$\sigma_*(x) \in \arg\max_{a\in[0,1]} Na \int_0^\infty w_*(x+s) - w_*(x)\mathrm{d}F(s) - rc(a,x). \tag{5}$$

The following result shows that, in order to maximise welfare, agents should exert less effort as the stock grows, and that any increment in x_t makes agents better-off. The proof is in the Appendix and is outlined below, before the example.

Theorem 1 (Social-welfare benchmark)**.** *There exists a symmetric profile σ_* of Markov strategies maximising welfare for all initial stocks $x_0 \geq 0$. Effort $\sigma_*(x_t)$ is decreasing in the stock x_t; it vanishes as $x_t \to \infty$ if*

$$\forall a > 0 \quad \lim_{x\to\infty} N\mu_F b'(x) - rc_1(a,x) < 0 \tag{6}$$

and does not vanish if

$$\exists a > 0 \quad \lim_{x\to\infty} N\mu_F b'(x) - rc_1(a,x) > 0. \tag{7}$$

Maximal welfare $w_*(x_0)$ is increasing and continuous in the initial stock x_0.[13]

The efficient effort schedule σ_* is symmetric since agents are identical and the cost of effort is convex (i.e. $c(a,x)$ is convex in a). In contrast, the analysis in Chapter 3 shows that (in the special case of linear cost) equilibrium welfare is maximised by an asymmetric profile when effort is unobservable.

The fact that effort σ_* is decreasing is standard in dynamic public-good games, and is generally due to concave payoffs. In this model, not only are payoffs $b(x_t)$

13. Unless $c(a,x)$ is strictly convex in a, σ_* need not be unique. For instance, in the example at the end of this section, both $a_t^i = 1$ and $a_t^i = 0$ are optimal at $x_t = N/(\lambda r\alpha)$. Moreover, the limits in (6) and (7) exist since b' is decreasing (since b is concave), and $c_1(a,x)$ is increasing in x.

concave, but the marginal cost of effort $c_1(a, x_t)$ increases with x_t, at an increasing rate (i.e. $c_{11}(a, x_t)$ is increasing in x_t). The latter force adds to the former and does not alter the qualitative features of the social-welfare benchmark. However, as shown in the next section, it plays an important role in the symmetric equilibrium.

Perhaps surprisingly, if the assumption that $c_{11}(a, x)$ is increasing in x is dropped, σ_* need not be decreasing. That is, it can be optimal to exert *more* effort as the stock grows, even though the marginal opportunity cost of effort increases (i.e. $c_{12}(a, x) \geq 0$).[14]

Condition (6) states that, if the stock is sufficiently large, the social value of an innovation is offset by the marginal cost of effort. The value of an innovation is approximately $N\mu_F\beta'(x)$. In particular, it is increasing in the average size μ_F of increments, but higher moments of their distribution do not play a role. This is because, since b is concave and increasing, it is approximately linear if the stock is large. The marginal cost $rc_1(a, x)$ is adjusted by the time discount factor r: this is because the cost of effort is incurred when the effort is exerted, whereas its benefits occur in the future (since innovations are generated with a lag) and, moreover, they are permanent. If σ_* reaches 0 at some $\hat{x} < \infty$, then $w_*(x) = b(x)$ for $x \geq \hat{x}$. Then, (5) implies that

$$N \int_0^\infty b(\hat{x} + s) - b(\hat{x}) \mathrm{d}F(s) = rc_1(0, \hat{x}). \tag{8}$$

The left-hand side of (8) is the social value of an innovation when the size of the stock is $\hat{x}$; the right-hand side is the marginal cost of exerting a small amount of effort. Conversely, if (8) admits a solution $\hat{x}$, then σ_* reaches 0 at the least $\hat{x}$ such that (8) holds.[15]

If neither (6) nor (7) hold then there exists $a' > 0$ such that (6) holds for all $a > a'$, and holds with '=' for all $a < a'$, effort σ_* may or may not vanish as x_t

14. For instance, this occurs if $F(s) = 1 - e^{-s}$, $u(a, x) = (1 + x)(1 - a) - \frac{1}{2}a^2)e^{-x}$, and $N/r = 3$. In this case, (8) implies that effort ceases at $\hat{x} = 2$. Then, (4) and (5) imply that, on $[0, 2]$, σ_* is given by

$$\sigma_*'(x) = \frac{(\sigma_*(x)/2 + 2r/N)\sigma_*(x) - (1 - xr/N)e^x}{\sigma_*(x) + r/N}.$$

Simulations show that σ_* is strictly concave on $[0, 2]$ and maximised around 0.8.

15. To see why this is the case, let $\hat{x}$ solve (8). Since b is concave, the LHS of (8) is decreasing in $\hat{x}$. Since $c_1(0, x)$ is increasing in x, it follows that (8) holds with '$\leq$' for all $x \geq \hat{x}$. Hence $\sigma_*(x) = 0$ solves (5) for all $x \geq \hat{x}$ if $w_* = b$. Since (4) is forward looking, it follows that $w_*(x) = b(x)$ for $x \geq \hat{x}$. Result follows.

diverges. For instance, given a parameter $\beta \in \{0,1\}$, if $b(x) = r\,(x - \beta e^{-x})$, and $c(a,x) = N\mu_F a$, then (6) holds with '=' for all $a \in [0,1]$. Yet, $\sigma_*(0) = 0$ if $\beta = 0$, and $\lim_x \sigma_*(x) = 1$ if $\beta = 1$.[16] However, either (6) or (7) must hold if $c(a,x)$ is strictly convex in a, since $c_{11}(a,x)$ is increasing in x.

Maximal welfare $w^*(x_0)$ is increasing because raising x_0 increases payoffs (since u is increasing in x) without altering the productivity of any effort level (as the arrival and size of innovations does not depend on x_t). Specifically, given a profile σ inducing welfare $w(x_0)$ in the game with initial stock x_0, agents may obtain welfare $w(x_0)$ or higher in any game with initial stock $x_0' > x_0$ by replicating the behaviour induced by σ: that is, by setting $a_t^i = \sigma(x_t - x_0' + x_0)$. Moreover, since average continuation payoffs when σ_* is played equal $w_*(x_t)$, any innovation in the stock is beneficial. We shall see that, despite this, welfare need not increase with x_0 *in equilibrium*, so that innovations may be detrimental.

Theorem 1 is proved by constructing the function w_* as a fixed point of a mapping P, defined as follows. For an arbitrary welfare function w, $Pw(x_0)$ is the maximal welfare achievable in the game with initial stock x_0, given that continuation welfare after the first innovation is given by $w(x_0 + s)$, where s is the size of the innovation. Attention is restricted to (Lipschitz) continuous w's that (a) lie between b and a function $\bar{w}$ that grows linearly, and (b) are increasing, but at a lower rate than that of b. A version of Tarski's theorem (Lemma 1) implies that the fixed point w_* exists. The fact that $w_*'(x) \leq b'(x)$ for all $x \geq 0$ is used to show that σ_* is decreasing.

Example. Figure 1 depicts effort and welfare in the social-welfare benchmark of the example introduced in Section 2. R&D effort σ_* ceases at $\hat{x} = N/(\lambda r \alpha)$. Moreover, since the cost of effort is linear, σ_* exhibits a bang-bang feature: $\sigma_*^i(x) = 1$ for $x < \hat{x}$, and $\sigma_*(x) = 0$ for $x > \hat{x}$. Welfare w_* is given by

$$w_*(x) = \begin{cases} \dfrac{\alpha N}{\lambda r} e^{\frac{r\lambda x - N/\alpha}{r+N}} + (1-\alpha)\left[x + \dfrac{N}{\lambda r}\right] & \text{if } x \leq N/(\lambda r \alpha) \\ x & \text{if } x \geq N/(\lambda r \alpha). \end{cases}$$

16. Indeed, if $\beta = 0$, then $\sigma_*(0) = 0$ since (8) holds at $\hat{x} = 0$. If $\beta = 1$ then, since $c(a,x)$ is linear in a, (5) implies that w_* is induced by an effort schedule σ_* such that $\sigma_*(x) = \mathbb{I}\{x < \hat{x}\}$, for some $\hat{x} \in [0,\infty]$. But (8) holds with '>' for all $x \geq 0$. Hence $\sigma_*(x) = 1$ for all $x \geq 0$.

Note that w_* is not differentiable at $x = N/(\lambda r \alpha)$. In particular, the smooth-pasting property does not hold in this setting, as the state x_t does not evolve continuously.

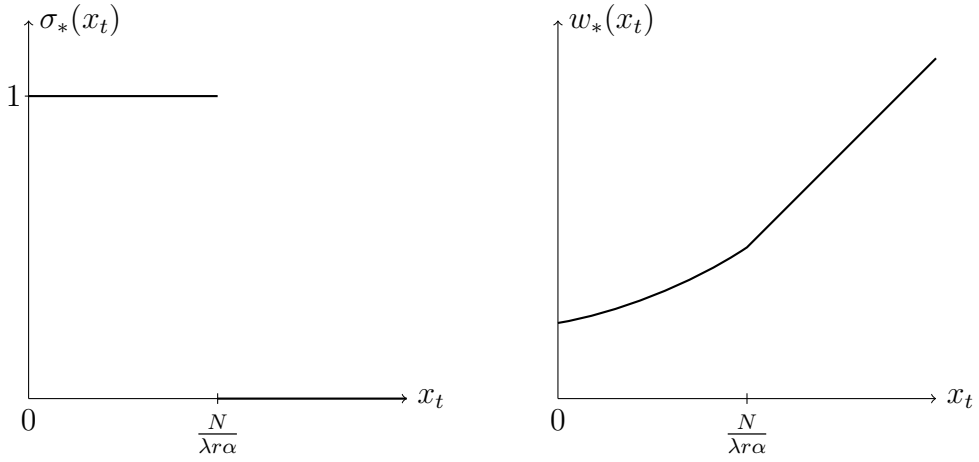


Figure 1: Effort (left) and welfare (right) in the social-welfare benchmark of the example, as functions of the stock x_t.

4 Symmetric Equilibrium

This section is devoted to the unique symmetric Markov-perfect equilibrium of the game. Among other things, I show that equilibrium welfare may drop after an innovation.

If the opponents of Agent i play Markov strategies $(\sigma_j)_{j\neq i}$, then there exists a Markov strategy σ^i that is a best-response to $(\sigma_j)_{j\neq i}$ in the game with initial stock x_0, for all $x_0 \geq 0$. From (2), the value function of Agent i solves the Bellman equation

$$rv^i(x) = \max_{a\in[0,1]} r[b(x) - c(a,x)] + \left[a + \textstyle\sum_{j\neq i} \sigma^j(x)\right] \displaystyle\int_0^\infty v^i(x+s) - v^i(x) \mathrm{d}F(s) \quad (9)$$

and her Markov best-responses are the strategies σ^i such that

$$\sigma^i(x) \in \arg\max_{a\in[0,1]} \int_0^\infty v^i(x+s) - v^i(x) \mathrm{d}F(s) - rc(a,x).^{17} \quad (10)$$

A profile σ of Markov strategies is a *Markov Perfect Equilibrium* (MPE) if it is a

17. The value function of Agent i is the supremum of initial payoffs that she can achieve, by playing any strategy (including non-Markov ones), when her opponents play $(\sigma_j)_{j\neq i}$.

Nash equilibrium of the game with initial stock x_0, for any $x_0 \geq 0$. The following result characterises the unique symmetric MPE of the game.

Theorem 2 (Symmetric MPE)**.** *Suppose that* (6) *holds. Then, there exists a unique symmetric MPE* σ_s*, and it is inefficient for some* x_0*. Effort* $\sigma_s(x_t)$ *is continuous and decreasing in the stock* x_t*, and lies below the benchmark* $\sigma_*(x_t)$*. Welfare* $w_*(x_0)$ *is continuous in the initial stock* x_0*.*

As the stock x_t increases, equilibrium effort $\sigma_s(x_t)$ decreases, and does so inefficiently fast in general. This feature is standard in dynamic public-good games, and is caused by *inter-temporal free-riding* (a term coined by Marx and Matthews 2000): since σ_s is decreasing, agents are reluctant to exert effort as this causes their opponents to exert less effort in the future. As a consequence, σ_s is not efficient in general.

Apart from the trivial case in which no effort is optimal, σ_s *is* efficient if the incentive to exert effort is very strong initially and is guaranteed to drop suddenly, so that $\sigma_s(x_t)$ and $\sigma_*(x_t)$ drop from 1 to 0 at the same time. For this to occur, the size of innovations has to be large (i.e. $F(s) = 0$ for some large $s > 0$).

The encouragement effect does not arise in σ_s. Indeed, if effort ceases at some cutoff $\hat{y} \in (0, \infty)$, then (10) implies that $\hat{y}$ solves

$$\int_0^\infty b\left(\hat{y} + s\right) - b\left(\hat{y}\right) \mathrm{d}F(s) = rc_1\left(0, \hat{y}\right). \tag{11}$$

Comparing (11) to (8) shows that effort ceases exactly when it stops in the single-agent problem; that is, there is no encouragement effect. We shall see that this does not hinge on the restriction to symmetric equilibria, nor to Markov strategies. Indeed, no perfect Bayesian equilibrium features the encouragement effect (c.f. Chapter 2, Lemma 1). This contrasts the findings of Keller, Rady, and Cripps (2005): in their model, the encouragement effect does not arise in the symmetric MPE, but it occurs in another, asymmetric, MPE. The difference between the equilibria is that, in the latter, agents take turns at exerting effort and switch infinitely many times as the state approaches the stopping threshold. That cannot occur in this model since agents have no control over the size of innovations.

If condition (6) fails, σ_s continues to exist, but may not be unique. As part of the proof of Theorem 2, I show that there exist equilibria $\underline{\sigma}_s$ and $\bar{\sigma}_s$, inducing welfare $\underline{w}_s$ and $\bar{w}_s$ such that, for any equilibrium σ_s, $\underline{\sigma}_s(x) \leq \sigma_s(x) \leq \bar{\sigma}_s(x)$, and $\underline{w}_s(x) \leq w_s(x) \leq \bar{w}_s(x)$.[18] That is, $\underline{\sigma}_s$ (resp. $\bar{\sigma}_s$) is the equilibrium in which least (resp. most) effort is exerted, and payoffs are lowest (highest), regardless of the value of the stock. This structure arises since, although the effort choices $\sigma^i(x)$ and $\sigma^j(x)$ of distinct agents at the same stock level x are strategic substitutes, their choices $\sigma^i(x)$ and $\sigma^j(x')$ at distinct levels $x' \neq x$ are strategic 'complements'. That is, if $x' > x$, the best-response $\sigma^i(x)$ is increasing in $\sigma^j(x')$.[19]

The argument for the existence and uniqueness of σ_s is similar to the argument used by Bolton and Harris (1999) for the analogous result.[20] First, I show that for any continuation welfare function w, and any initial stock $x_0 \geq 0$, there is a unique effort level $Qw(x_0)$ such that, if opponents exert constant effort $Qw(x_0)$ until the stock rises, and continuation payoffs after the innovation are given by $w(x_0 + s)$ (where s is the size of the increment), then exerting effort $Qw(x_0)$ is a best-response (c.f. Claim 3 in Step 1 of the proof). The uniqueness of $Qw(x_0)$ is due to the fact that, as noted above, effort choices by distinct agents at the same stock level are strategic substitutes.

Equilibrium welfare w_s is the fixed point of the mapping P_s such that, for an arbitrary welfare function w and $x \geq 0$, $P_s w(x_0)$ is welfare when effort $Qw(x_0)$ is exerted until the first innovation, and continuation payoffs are given by w after the innovation occurs. An important condition for the existence of the fixed point is that the map P_s be increasing. That is, if $w_1(x) \geq w_2(x)$ for all $x \geq 0$, the $P_s w_1(x) \geq P_s w_2(x)$ for all $x \geq 0$. This holds since, if $w_1(x) \geq w_2(x)$ for all $x \geq 0$, then the benefit of an innovation is higher under w_1 than under w_2, hence

18. See Step 1 in the proof of Theorem 2.

19. This can be seen by rephrasing the Bellman equation (9) for the value function of Player i as

$$v^i(x) = \max_{a \in [0,1]} \frac{r[b(x) - c(a,x)]}{r + a + \sum_{j \neq i} \sigma^j(x)} + \frac{a + \sum_{j \neq i} \sigma^j(x)}{r + a + \sum_{j \neq i} \sigma^j(x)} \int_0^\infty v^i(x+s) \mathrm{d}F(s).$$

Clearly, the maximand is decreasing in $\sigma^j(x)$, and increasing in $v^i(x')$ for all $x' > x$. If $\sigma^j(x')$ increases, then so does $v^i(x')$, hence the best-response $\sigma^i(x)$ does, too.

20. Given that difference in the settings, and the generality of Theorem 2, the current proof is not a straightforward adaptation of theirs.

$Qw_1(x) \geq Qw_2(x)$ for all $x \geq 0$ (c.f. Part 1 of Claim 5). In turn, this implies that $P_s w_1(x) \geq P_s w_2(x)$, since equilibrium effort $Qw(x)$ is inefficiently low compared to $\sigma_*(x)$, hence an increase in effort corresponds with an increase in efficiency (c.f. Part 2 of Claim 5).

As noted above, this argument shows the existence of equilibria $\underline{\sigma}_s$, and $\bar{\sigma}_s$, which minimise (resp. maximise) effort and welfare among equilibria at all stock levels. This structure is crucial for the proof of uniqueness. Since $\bar{\sigma}_s(x) \geq \underline{\sigma}_s(x)$ for all $x \geq 0$, then (10) implies that

$$\int_0^\infty \bar{w}_s(x+s) - \bar{w}_s(x) \mathrm{d}F(s) \geq \int_0^\infty \underline{w}_s(x+s) - \underline{w}_s(x) \mathrm{d}F(s).$$

That is, the expected value of an innovation is higher when $\bar{\sigma}_s$ is played than when $\underline{\sigma}_s$ is played. Rearranging gives:

$$\int_0^\infty \bar{w}_s(x+s) - \underline{w}_s(x+s) \mathrm{d}F(s) \geq \bar{w}_s(x) - \underline{w}_s(x).$$

Hence, roughly speaking, the distance between the continuation payoffs increases as the stock rises. But, since maximal equilibrium effort $\bar{\sigma}_s$ lies below the efficient schedule σ_*, and the latter vanishes as the stock diverges (since (6) holds), so do $\bar{\sigma}_s$ and $\underline{\sigma}_s$. Hence, in the limit as the stock diverges, both $\bar{w}_s$ and $\underline{w}_s$ approximate b. This implies that $\bar{w}_s = \underline{w}_s$, hence the equilibrium is unique.

Example. Figure 2 depicts R&D effort and welfare in the symmetric MPE of the example. R&D ceases at $\hat{y} = 1/(\lambda r \alpha)$, and is given by

$$\sigma_s(x) = \begin{cases} 1 & \text{if } x \leq \tilde{y} \\ \dfrac{1}{N-1}\left[r\alpha\left(\dfrac{\lambda}{2}x - 1\right) + \dfrac{1}{\lambda x}\left(\dfrac{1}{2r\alpha} + 1\right) - 1\right] & \text{if } \tilde{y} \leq x \leq 1/(\lambda r \alpha) \\ 0 & \text{if } x \geq 1/(\lambda r \alpha) \end{cases}$$

where N is the number of agents and $\tilde{y}$ is such that σ_s is continuous at $\tilde{y}$. In

particular, $\tilde{y}$ is decreasing in N with $\lim_{N\to\infty} \tilde{y} = 0$. Welfare is given by

$$w_s(x) = \begin{cases} N\left(\alpha\tilde{y} - \dfrac{1-\alpha}{\lambda r}\right) e^{\frac{r\lambda}{r+N}(x-\tilde{y})} + (1-\alpha)\left[x + \dfrac{N}{\lambda r}\right] & \text{if } x \leq \tilde{y} \\ r\alpha x\left(\dfrac{\lambda}{2}x - 1\right) + \dfrac{1}{\lambda}\left(\dfrac{1}{2r\alpha} + 1\right) & \text{if } \tilde{y} \leq x \leq 1/(\lambda r\alpha) \\ x & \text{if } x \geq 1/(\lambda r\alpha). \end{cases}$$

Note that, for $x \geq \tilde{y}$, welfare $w_s(x)$ does not depend on N. That is, as the number of agents increases, each decreases her individual effort (hindering efficiency) by an amount that perfectly offsets the gains from the extra agents. This is due to the linear cost of effort (i.e. $c(a, x) = a\alpha x$).

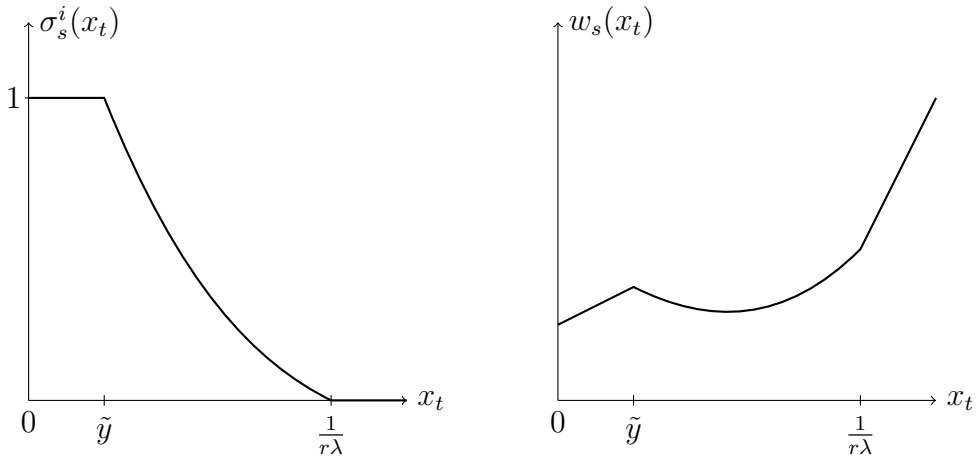


Figure 2: Effort (left) and welfare (right) in the symmetric Markov equilibrium of the example, as functions of the stock x_t. Welfare is non-monotone whenever $N \geq 1 + (1/r - r)/2$.

4.1 Non-Monotonicity

In this section, I derive conditions guaranteeing that welfare w_s in the symmetric MPE is increasing, and conditions under which it is non-monotone, and provide some intuition.

Remember that an agent exerting effort a_t receives a flow of payoffs $b(x_t) - c(a_t, x_t)$, where x_t is the current stock. In particular, $c_1(a_t, x_t)$ denotes the marginal cost of effort a_t at stock level x_t. The next result gives sufficient conditions for monotonicity and for non-monotonicity. The first part is proved in the Appendix,

and the second is proved below.

Proposition 1 (Non-monotonicity)**.** *Suppose that* (6) *holds. If* $c(a,x) = m(a)$*, then welfare* w_s *in the symmetric MPE is increasing in the initial stock* x_0*. However, if (i)* $c_1(0,0) = 0$ *and* $c_1(0,x) > 0$ *for* $x > 0$*, (ii) the distribution* F *of the size of innovations has unbounded support, and (iii) there are sufficiently many agents, then* w_s *is non-monotone.*

If $c(a,x) = c(a)$, then the opportunity cost of effort does not increase with the stock. In this case, even though less effort is exerted as the stock increases (σ_s is decreasing), so that further innovations are delayed, the extra benefits overcome this delay, raising continuation payoffs. This is a standard feature of symmetric equilibria of dynamic public-good games.

If $c_1(a,0) = 0$, then the incentive to exert at least effort a becomes arbitrarily large as the stock shrinks to 0.[21] In contrast, if $c_1(0,x) < 0$ for $x > 0$, then the incentive to exert effort is bounded if the stock is strictly positive. Both conditions are satisfied in the example since $c_1(a,x) = x$ for all $a \in [0,1]$ and $x \geq 0$.

Proof of the second part of Proposition 1. Suppose that there are sufficiently many agents (i.e. if N is sufficiently large). If $c_1(0,x) < 0$ for some $x > 0$ then, (6) implies that (11) is satisfied by some $\hat{y} < \infty$. Hence, no effort is exerted above $\hat{y}$. Moreover, since there is no encouragement effect, $\hat{y}$ does not grow with N. Hence, for any fixed stock value x, the benefit of an innovation remains bounded as N grows. Therefore, if $x > 0$ so that effort is costly (i.e, c is strictly increasing in a), $\sigma_s(x)$ shrinks to 0 as N grows. This is because effort choices in the same period are strategic substitutes. In particular, $\sigma(x) < 1$ for sufficiently large N. Then, the first-order condition for $\sigma(x)$ gives

$$\int_0^\infty w_s(x+s)\mathrm{d}F(s) - w_s(x) = rc_1(\sigma_s(x), x). \tag{12}$$

This expression is similar to (11); the left-hand side is the expected (individual) value of an innovation given that σ will be played after it occurs, and the right-hand side is the marginal cost of producing an innovation. Since $c_1(0,0) = 0$ and

21. Notice that this condition is strong, however it guarantees non-monotonicity regardless of the values of the rest of the parameters. Moreover, by continuity, increasing $c_1(a,0)$ by a sufficiently small amount does not restore monotonicity as long as the rest of the parameters remain fixed.

$\sigma_s(x)$ shrinks to 0 as N grows then, if $x > 0$ is sufficiently small and, given x, N is sufficiently large,

$$w_s(0) \approx \int_0^\infty w_s(s)\mathrm{d}F(s). \tag{13}$$

Intuitively, since the incentive to exert some effort is large when the stock is low, aggregate effort is high when N is large, and the stock increases immediately. If w were increasing, then (13) would hold with '$\leq$' and, since F has unbounded support, equality would hold if and only if w is constant. However, $w_s(x) = x$ for $x > x'$ since no effort is exerted. Hence w_s is not increasing. □

To gain an insight into the mechanism producing the non-monotonicity, note that for any two $x' > x > 0$ such that $w_s(x) > w_s(x')$, some effort must be exerted at stock x, so that (12) holds with '$\geq$'. Then

$$\int_{x'-x}^\infty w_s(x+s) - w_s(x)\mathrm{d}F(s) \geq rc_1(\sigma_s(x), x) - \int_0^{x'-x} w_s(x+s) - w_s(x)\mathrm{d}F(s).$$

Each integrand is the benefit from an innovation of size s in the stock when its value is x. On the left-hand side, the integral covers all sizes s increasing the stock above x', and it covers the rest on the right-hand side. Assuming that w_s is quasi-convex above x (such as in the example presented above for $x \geq \tilde{y}$), the integral on the right-hand side is negative. This means that, at x, the benefit from the prospect of overtaking x' with the next innovation exceeds the cost of inducing it. Hence, not only is there a strictly positive probability of leapfrogging any larger stock sizes inducing lower continuation payoffs – the likelihood and the benefit of this event are sufficiently large to justify the costs.

Since the non-monotonicity is caused by an over-taking phenomenon, it is unlikely to arise in a model which does not feature discrete increments in the stock. In fact, the stochastic size of jumps is also crucial: if F puts all mass on a single value then, although w_s remains non-monotone (Theorem 2 still applies), $w_s(x_t)$ is (almost surely) increasing in t, regardless of the initial stock x_0.

The fact that welfare is monotone in the social-welfare benchmark highlights the importance of incentives. Equilibrium welfare may drop as the stock increases because effort drops inefficiently fast. This is due to inter-temporal free-riding,

discussed above.

Example. When is welfare w_s increasing in the example? Note that w_s is increasing on $[0, \tilde{y}]$ and convex on $[\tilde{y}, \infty)$. Hence, it is increasing if and only if $\lim_{\epsilon \downarrow 0} w_s'(\tilde{y} + \epsilon) \geq 0$. This is equivalent to

$$N \leq 1 + \frac{1}{2}\left(\frac{1}{r} - r\right).$$

In particular, w_s is increasing if and only if, keeping other parameters fixed, agents are sufficiently patient. Interestingly, the condition does not depend on λ.

How likely is $w(x_t)$ to drop? Note that, if N is sufficiently large, then w_s is minimised at $1/\lambda$. If $x_0 \in (0, 1/\lambda)$ and N is large enough so that $\tilde{y} \leq x_0$, the probability that welfare drops after the first innovation is $1 - e^{-2(1-\lambda x_0)}$. The largest drop occurs when x_0 increases to $1/\lambda$. If $x_0 = 0$, in the limit as N is large

$$\frac{w_s(1/\lambda)}{w_s(0)} = 1 - \frac{2r^2\alpha^2}{1 + 2r\alpha}$$

The RHS equals $1/3$ when $r\alpha = 1$.

5 Conclusion

I analysed a dynamic game of private provision of a public good that is produced in lumps of random size. I showed that, under realistic assumptions, continuation payoffs may drop after an increment in the stock of the good in the unique symmetric Markov equilibrium of the game. This is because, even though a larger stock enhances the agents' flow payoffs, further production is delayed, and the future losses may offset the present gains. In the context of an R&D parternship, these results imply that marginal innovations of a shared technology may be detrimental, as they reduce the incentive to further improve upon them.

Since innovations may be detrimental for the agents that induce them, if agents could selectively conceal the innovations they obtain, they are likely to have an incentive to do so. However, it is not clear whether their opponents would benefit, too. This idea is explored in Chapter 3.

References

Admati, A., and M. Perry. 1991. "Joint projects without commitment." *The Review of Economic Studies* 58 (2): 259–276.

Allen, Robert C. 1983. "Collective invention." *Journal of economic behavior & organization* 4 (1): 1–24.

Battaglini, M, S Nunnari, and T R Palfrey. 2014. "Dynamic free riding with irreversible investments." *American Economic Review* 104 (9): 2858–71.

Bimpikis, Kostas, Kimon Drakopoulos, and Shayan Ehsani. 2018. *Disclosing Information in Strategic Experimentation.* Working Paper 3635. Stanford Graduate School of Business.

Bliss, Christopher, and Barry Nalebuff. 1984. "Dragon-slaying and ballroom dancing: The private supply of a public good." *Journal of Public Economics* 25 (1-2): 1–12.

Bolton, P., and C. Harris. 1999. "Strategic experimentation." *Econometrica* 67 (2): 349–374.

Bonatti, Alessandro, and Johannes Hörner. 2011. "Collaborating." *American Economic Review* 101 (2): 632–63.

Cowan, Robin, and Nicolas Jonard. 2003. "The dynamics of collective invention." *Journal of Economic Behavior & Organization* 52 (4): 513–532.

Dasgupta, Partha. 1988. "Patents, priority and imitation or, the economics of races and waiting games." *The Economic Journal* 98 (389): 66–80.

Davey, Brian A, and Hilary A Priestley. 2002. *Introduction to lattices and order.* Cambridge university press.

Fershtman, C., and S. Nitzan. 1991. "Dynamic voluntary provision of public goods." *European Economic Review* 35 (5): 1057–1067.

Geroski, Paul. 1995. "Markets for technology: knowledge, innovation and appropriability." Chap. 4 in *Handbook of the Economics Innovation and Technological Change,* edited by Paul Stoneman. Blackwell.

Gradstein, Mark. 1992. "Time dynamics and incomplete information in the private provision of public goods." *Journal of political Economy* 100 (3): 581–597.

Katz, Michael L. 1986. "An analysis of cooperative research and development." *The RAND Journal of Economics:* 527–543.

Keller, G., and S. Rady. 2010. "Strategic experimentation with Poisson bandits." *Theoretical Economics* 5 (2): 275–311.

Keller, G., S. Rady, and M. Cripps. 2005. "Strategic experimentation with exponential bandits." *Econometrica* 73 (1): 39–68.

Marx, L., and S. Matthews. 2000. "Dynamic voluntary contribution to a public project." *The Review of Economic Studies* 67 (2): 327–358.

Meyer, Peter B. 2003. *Episodes of collective invention.* Working Paper 368. U.S. Bureau of Labour Statistics.

Nuvolari, Alessandro. 2004. "Collective invention during the British Industrial Revolution: the case of the Cornish pumping engine." *Cambridge Journal of Economics* 28 (3): 347–363.

Powell, Walter W, and Eric Giannella. 2010. "Collective invention and inventor networks." In *Handbook of the Economics of Innovation,* 1:575–605. Elsevier.

Powell, Walter W, Kenneth W Koput, and Laurel Smith-Doerr. 1996. "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology." *Administrative science quarterly:* 116–145.

Taylor, Curtis R. 1995. "Digging for golden carrots: An analysis of research tournaments." *The American Economic Review:* 872–890.

Wang, Wen-Kai, and Christian-Oliver Ewald. 2010. "Dynamic voluntary provision of public goods with uncertainty: a stochastic differential game model." *Decisions in Economics and Finance* 33 (2).

6 Appendix

6.1 Histories, strategies, and payoffs

In this section, I define rigorously histories, strategies, and payoffs of the game.

Histories. A *history up to time 0* is $x_0 \in [0, \infty)$. For $M \in \mathbb{N}$, a *history up to (innovation) time* M is

$$h_M = \left(x_0, (t_m, x_m, i_m)_{m=1}^M\right)$$

where $x_0 \geq 0$, $(t_m, x_m)_{m=1}^M \subset [0, \infty) \times [x_0, \infty)$ is increasing and, for all $m \leq M$, $i_m \in \{1, \ldots, N\}$. The history h_m records publicly-available information up to game-time t_M, for cases in which M innovations occur prior to time t_M. Specifically, the m^{th} innovation occurrs at time t_m, raises the stock to level x_m, and is generated by Player i_m. For $M \in \mathbb{N} \cup \{0\}$, let H_M be the set of all histories up to time M. Let $\mathcal{H}_M$ be the Borel σ-algebra of H_M when the latter is viewed as a subset of $\mathbb{R}^{3M+1}$.

Strategies. Let A_0 be the set of Borel measurable maps $a : [0, \infty) \to [0, 1]$, and let A be the quotient space in which any two $a \in A$ that are equal almost everywhere are identified. Let $\alpha : A_0^2 \to [0, \infty)$ be given by

$$\alpha(a, a') = \int_0^\infty |a(t) - a'(t)| e^{-t} \mathrm{d}t.$$

Note that if $a, a', b, b' \in A$ are such that $a \overset{\text{a.e.}}{=} b$ and $a' \overset{\text{a.e.}}{=} b'$, then $\alpha(a, a') = \alpha(b, b')$. Then, α may be viewed as a map from A^2 to $[0, \infty)$. Moreover, α is a metric on A. Let $\mathcal{A}$ be the Borel σ-algebra induced on A by α.

A (pure) *strategy* is a family $(\sigma_m)_{m=0}^{\infty}$ such that, for all $m \in \mathbb{N}_0$, σ_m is a measurable map from $(H_m, \mathcal{H}_m)$ to $(A, \mathcal{A})$. If at least $m \geq 0$ innovations occur, σ_m pins down behaviour between the time t_m of the m^{th} innovation and the time t_{m+1} of the $(m+1)^{\text{th}}$ innovation, or after time t_m if exactly m innovations take place. Namely, after a history h_m, σ prescribes to exert effort $[\sigma_m(h_m)](t - t_m)$ at all $t \in [t_m, t_{m+1})$ if the $(m+1)^{\text{th}}$ innovation occurs at time t_{m+1}, and at all $t \geq t_m$ if precisely m innovations take place.

Let Σ be the set of all strategies, and denote strategy profiles by $\sigma = (\sigma^i)_{i=1}^N \in \Sigma^N$.

6.2 Notation

The following piece of notation is used throughout the appendix. Given a set X and $f, g : X \to \mathbb{R}$, write '$f \geq g$' for '$f(x) \geq g(x)$ for all $x \in X$'.

A function $w : [0, \infty) \to \mathbb{R}$ is *absolutely continuous* if w is differentiable almost everywhere and, for all $x > 0$,

$$w(x) = w(0) + \int_0^x w'(s)\mathrm{d}s$$

for some function w' that is integrable and such that $w'(x)$ equals the derivative of w for all $x > 0$ at which w is differentiable. Throughout the appendix, if a function $w : [0, \infty) \to \mathbb{R}$ is absolutely continuous, statements about its derivative $w'(x)$ implicitly assume that attention is restricted to x's such that $w'(x)$ exists.

A function $w : [0, \infty) \to \mathbb{R}$ is *Lipschitz continuous with constant* $k > 0$ if, for all $x_1, x_2 \geq 0$,

$$|w(x_2) - w(x_1)| < k|x_2 - x_1|.$$

Lipschitz continuity implies absolute continuity.

6.3 Proofs

Let C be the set of measurable functions $w : [0, \infty) \to \mathbb{R}$ such that, for some $a, b, c \in \mathbb{R}$ with $a \leq c$ and $b \geq 0$, $w(x) \in [a, bx + c]$ for all $x \geq 0$. For $w \in C$, let

$$Lw(x) \equiv \int_0^\infty w(x + s)\mathrm{d}F(s).^{22} \tag{14}$$

Let $T : [0, 1] \times [0, N] \times [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be given by

$$T\left(a, a', x, l\right) = \frac{r[b(x) - c(a, x)] + a'l}{r + a'}, \tag{15}$$

and let $\bar{T} : [0, 1]^2 \times [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be given by

$$\bar{T}\left(a, x, l\right) = T(a, Na, x, l). \tag{16}$$

From (3), $T(a, a', x, l)$ are the payoffs of an agent exerting constant effort a until the next innovation, given that aggregate effort exerted is also constant and equals a', and her expected continuation payoffs after the first innovation equal l. The next two results are used in the proofs of Theorem 1 and 2.

Lemma 1 (Fixed-Point Theorem)**.** *Let $C' \subseteq C$ be such that any chain in C' has a join and a meet with respect to the point-wise order '$\leq$'. Let $P : C' \to C'$ be order-preserving. Then P has least and greatest fixed-points in C'.*[23]

Proof. Note that $(C, \leq)$ is a partially-ordered set. That is, for any $w, v, u \in C$, $w \leq w$, $w \leq v$ and $v \leq u$ imply $w \leq u$, and $w \leq v$ and $v \leq w$ imply $w = v$. Since $C' \subseteq C$, $(C', \leq)$ is a partially-ordered set, too. Then, the result is a simple consequence of the 'CPO fixed-point theorem II' of Davey and Priestley (2002), page 187. From their Theorem 8.11, on page 179, $(C', \leq)$ is a CPO. Then, the CPO fixed-point theorem II implies that P has a least fixed point. The same argument applied to $(C', \geq)$, where '$w \geq v$' if and only if $v \leq w$ for all $v, w \in C'$,

22. Note that the integral exists since w is bounded below by a, and it takes a finite value since it lower than $b(x + \mu_F)$, and $\mu_F < \infty$.

23. Recall that, for $w, v \in C$, $w \leq v$ whenever $w(x) \leq v(x)$ for all $x \geq 0$. A chain in C' is a (possibly-empty) subset $A \subseteq C'$ such that, for any $w, v \in A$, $w \leq v$ or $v \leq w$; P is order preserving if, for all $w, v \in C'$ such that $w \leq v$, $P(w) \leq P(v)$.

implies the existence of a greatest fixed point. □

Lemma 2. *For $\underline{v}, \bar{v}, \tilde{v}, \hat{v} \in C$ such that $\underline{v} \leq \bar{v}$, and $\tilde{v} \leq \hat{v}$, let C' be the set of absolutely continuous $w \in C$ such that $\underline{v} \leq w \leq \bar{v}$ and $\tilde{v} \leq w' \leq \hat{v}$. Then, any chain in C' has a join and a meet (w.r.t. the point-wise order).*

Proof. Fix a chain $A \subseteq C'$. I construct the join. The construction of the meet is analogous. Let $\hat{w}$ be such that, for all $x \geq 0$,

$$\hat{w}(x) = \sup\{w(x) : w \in A\}.$$

Note that $\hat{w} \leq \bar{v}$ since $w \leq \bar{v}$ for all $w \in A$. Then $\hat{w}$ is well-defined. Moreover, it is clear that $\hat{w}$ is a join of A (in C). Hence, it remains to prove that $\hat{w} \in C'$. Since any $w \in A$ is measurable (as $w \in C$), then so is $\hat{w}$. Then, since $w \leq \hat{w} \leq \bar{v}$, and $w, v \in C$, it follows that $\hat{w} \in C$. Moreover, since $w \in A$, $w \geq \underline{v}$. Hence $\underline{v} \leq \hat{w} \leq \bar{v}$. It remains to show that $\hat{w}$ is absolutely continuous with $\tilde{v} \leq w' \leq \hat{v}$.

Claim 1. For any $x_2 \geq x_1 \geq 0$

$$\int_{x_1}^{x_2} \tilde{v}(x)\mathrm{d}x \leq \hat{w}(x_2) - \hat{w}(x_1) \leq \int_{x_1}^{x_2} \hat{v}(x)\mathrm{d}x. \tag{17}$$

Note that, since $\tilde{v}, \hat{v} \in C$, $\tilde{v}$ is bounded below, and $\hat{v}$ grows at most linearly. Hence, if the claim holds, $\hat{w}$ is absolutely continuous. Then, $\tilde{v} \leq \hat{w}' \leq \hat{v}$ follows by differentiating (17). Hence, it suffices to prove the claim.

I prove the right-most inequality. The argument for the left-most one is analogous. By definition of $\hat{w}$, for $i = 1, 2$, there exists a sequence $(w_n^i)_n \subseteq A$ such that $w_n^i(x_i) \to \hat{w}(x_i)$. Fix $n \in \mathbb{N}$. Since A is totally-ordered, there exists $w_n \in \{w_n^1, w_n^2\}$ such that $w_n^i \leq w_n$ for $i = 1, 2$. In particular, $w_n^i(x_i) \leq w_n(x_i)$ for all $i = 1, 2$. Since $w_n^i \leq \hat{w}$ for $i = 1, 2$, then $w_n(x_i) \in [w_n^i(x_i), \hat{w}(x_i)]$. Since n is arbitrary, and $\lim_n w_n^i(x_i) = \hat{w}(x_i)$, then $\lim_n w_n(x_i) = \hat{w}(x_i)$. Moreover, for each n, $w_n \in A$ so $w_n(x_2) - w_n(x_1) \leq \int_{x_1}^{x_2} \hat{v}(x)\mathrm{d}x$. Then, $\hat{w}(x_2) - \hat{w}(x_1) = \lim_n w_n(x_2) - w_n(x_1) \leq \int_{x_1}^{x_2} \hat{v}(x)\mathrm{d}x$. Result follows. □

6.4 Proofs for Section 3

Let $\bar{w} : [0, \infty) \to \mathbb{R}$ be such that

$$\bar{w}(x) = b(0) + \frac{b'(0)\mu_F}{r} + b'(0)x. \tag{18}$$

Lemma 3. *Given $w \in C$ such that $w \leq \bar{w}$, $T[a, a', x, Lw(x)] \leq \bar{w}(x)$ for all $a \in [0, 1]$, $a' \in [0, N]$, and $x \geq 0$, where L was defined in (14) and T in (15).*

Proof. Fix $x \geq 0$ and $w \in C$. Since $w \leq \bar{w}$, $Lw(x) \leq L\bar{w}(x)$. Moreover, for $a \in [0, 1]$,

$$\begin{aligned} L\bar{w}(x) &= b(0) + \frac{b'(0)\mu_F}{r} + b'(0)(x + \mu_F) \\ &\geq b(0) + b'(0)x \qquad (19) \\ &\geq b(x) \qquad (20) \\ &\geq b(x) - c(a, x) \end{aligned}$$

where (19) holds since $\mu_F, b'(0) \geq 0$ and (20) since b is concave. Then, for all $a' \in [0, N]$

$$\begin{aligned} T\left[a, a', x, Lw(x)\right] &\leq T\left[a, a', x, L\bar{w}(x)\right] \\ &\leq T[a, N, x, L\bar{w}(x)] \qquad (21) \\ &\leq \frac{r[b(0) + b'(0)x] + NL\bar{w}(x)}{r + N} \qquad (22) \\ &= \frac{r[b(0) + b'(0)x] + N[b(0) + (b'(0)\mu_F)/r + b'(0)(x + \mu_F)]}{r + N} \\ &= \bar{w}(x) \end{aligned}$$

where (21) holds since, from above, $L\bar{w}(x) \geq b(x) - c(a, x)$, and (22) holds since, from above, $b(x) - c(a, x) \leq b(0) + xb'(0)$. □

For $w \in C$ and $x \geq 0$, let

$$Pw(x) = \max_{a \in [0,1]} \bar{T}[a, x, Lw(x)] \tag{23}$$

where $\bar{T}$ is defined in (16).

Lemma 4. *Let $w_* \in C$. Then, for any $x \geq 0$, $w_*(x)$ solves (4) if and only if $w_*(x) = Pw_*(x)$.*

Proof. Note that

$$\begin{aligned} w_*(x) &= Pw_*(x) \\ &= \max_{a\in[0,1]} Lw_*(x) + \frac{r[b(x) - c(a,x) - Lw_*(x)]}{r + Na} \\ \Leftrightarrow \quad 0 &= \max_{a\in[0,1]} Lw_*(x) + \frac{r[b(x) - c(a,x) - Lw_*(x)]}{r + Na} - w_*(x) \\ \Leftrightarrow \quad 0 &= \max_{a\in[0,1]} NaLw_*(x) + r[b(x) - c(a,x)] - w_*(x)(r + Na). \end{aligned}$$

The last equation is equivalent to (4). Result follows. □

Proof of Theorem 1. It suffices to find some continuous and increasing $w_* \in C$ solving (4), as well as a symmetric Markov profile σ_* solving (5), such that σ_* is decreasing with $\lim_{x\to\infty} \sigma_*(x) = (>)\ 0$ if (6) holds (resp. if (7) holds). This is done in four steps.

Step 1. *(4) admits a continuous and increasing solution $w_* \in C$.* From Lemma 4, any fixed-point $w_* \in C$ of P solves (4). Then, it suffices to show that P has a continuous and increasing fixed point $w_* \in C$. Let C_* be the subset of $w \in C$ such that (i) $b \leq w \leq \bar{w}$, where $\bar{w}$ is defined in (18), and (ii) w is absolutely continuous and $0 \leq w' \leq b'$.[24] Note that any $w \in C_*$ is continuous and increasing. Then, it remains to show that P has a fixed-point in C_*.

Note that $b' \in C$ since b is increasing and concave. Then, from Lemma 2, any chain in C_* has a join and a meet. Moreover, P is clearly order-preserving. Then, from Lemma 1, it remains to show that P maps C_* to C_*.

Fix $w \in C_*$. Clearly, $Lw \in C$. Then $Pw \in C$ since u is increasing and concave in x. Fix $x \geq 0$ and note that

$$\begin{aligned} Pw(x) &\geq \bar{T}w[0, x, Lw(x)] \\ &= b(x). \end{aligned}$$

24. Absolute continuity is defined in Section 6.2.

Moreover, from Lemma 3, $Pw(x) \leq \bar{w}(x)$. Then, it remains to show that Pw is absolutely continuous and $0 \leq (Pw)' \leq b'$. Since $w \in C$ and w is absolutely continuous with $0 \leq w' \leq b'$, and b is increasing and concave, then w is Lipschitz with constant $b'(0)$.[25] Then, $Lw(x)$ is too. Moreover, since b and c are continuously differentiable with bounded derivative, the function $P_0(x,l) = \max_{a\in[0,1]} \bar{T}(a,x,l)$ is, too. Note that $Pw = P_0[x, Lw(x)]$. Then, Pw is Lipschitz, hence it is absolutely continuous. Then, it remains to show that $0 \leq (Pw)' \leq b'$. Fix x, and note that

$$\begin{aligned}(Lw)'(x) &= \int_0^\infty w'(x+s)\mathrm{d}F(s) \\ &\leq \int_0^\infty b'(x+s)\mathrm{d}F(s) \qquad (24)\\ &\leq b'(x) \qquad (25)\end{aligned}$$

where (24) holds since $w' \leq b'$, and (25) holds since b is concave. Moreover, by a similar reasoning, $Lw'(x) \geq 0$. Let $a \in \arg\max_{a'\in[0,1]} \bar{T}w\,(a',x)$. Then, by the enveloppe theorem,

$$\begin{aligned}(Pw)'(x) &= \frac{r[b'(x) - c_2(a,x)] + Na(Lw)'(x)}{r+Na} \qquad (26)\\ &\leq \frac{r[b'(x) - c_2(a,x)] + Nab'(x)}{r+Na} \qquad (27)\\ &\leq b'(x) \qquad (28)\end{aligned}$$

where (27) follows from (25), and (28) holds since c is increasing in x. Moreover, since $b(x) - c(a,x)$ is increasing in x and $(Lw)'(x) \geq 0$, it follows that $(Pw)'(x) \geq 0$. This proves that $Pw \in C_*$, and completes the construction of w_*.

Step 2. *There exists a symmetric and decreasing strategy profile σ_* solving (5).* Since $c(a,x)$ is continuous in a, the set

$$A_*(x) = \arg\max_{a\in[0,1]} Na[Lw_*(x) - w_*(x)] - rc(a,x)$$

is non-empty and closed, hence it has a maximum. Let σ_* be the symmetric profile such that $\sigma_*(x) = \max A_*(x)$ for all $x \geq 0$. Then σ_* solves (5).

25. That's because, for any $x_1, x_2 \geq 0$, $|w(x_2) - w(x_1)| = |\int_{x_1}^{x_2} w'(x)\mathrm{d}x| \leq |\int_{x_1}^{x_2} b'(x)\mathrm{d}x| \leq b'(0)|x_2 - x_1|$. Lipschitz continuity is defined in Section 6.2.

It remains to show that σ_* is decreasing. From Lemma 4,

$$\sigma_*^i(x) \in \arg\max_{a\in[0,1]} \bar{T}[a, x, Lw_*(x)]$$

for all $x \geq 0$. Note that $\bar{T}[a, x, Lw_*(x)]$ has a continuous partial derivative with respect to a for all $x \geq 0$, and $\bar{T}_1[a, x, Lw_*(x)]$ has the same sign as

$$g(a, x) = Lw_*(x) - b(x) + c(a, x) - (r/N + a)c_1(a, x).$$

Hence, it suffices to show that g is decreasing in a and x. To show that g is decreasing in a, fix $x \geq 0$ and note that, since c is twice continuously differentiable, $g(a, x)$ is continuously differentiable in a with

$$g_1(a, x) = -(r/N + a)c_{11}(a, x) \leq 0$$

where the inequality holds since c is convex in a. Hence g is decreasing in a.

To show that g is decreasing in x, fix $a \in [0, 1]$ and note that w_* is absolutely continuous (since $w_* \in C_*$) and b and c are twice continuously differentiable in x. Then $g(a, x)$ is absolutely continuous in x and

$$g_2(a, x) = (Lw_*)'(x) - b'(x) + c_2(a, x) - (r/N + a)c_{12}(a, x)$$

Since c_1 is twice continuously differentiable and $c_{11}(a, x)$ is increasing in x, then $c_{12}(a, x)$ is increasing in a, so $g_2(a, x)$ is decreasing in a, too. Then

$$\begin{aligned} g_2(a, x) &\leq g_2(0, x) \\ &= (Lw_*)'(x) - b'(x) - (r/N)c_{12}(0, x) \qquad (29) \\ &\leq (Lw_*)'(x) - b'(x) \qquad (30) \end{aligned}$$

where (29) holds since $c(0, x) = 0$, and (30) holds since c_1 is increasing in x. Then, (25) implies that $g_2(a, x) \leq 0$. Hence g is decreasing in x. This shows that σ_* is decreasing.

Step 3. *If* (6) *holds, then* $\lim_x \sigma_*(x) = 0$. Given that σ_* solves (5), it suffices to

show that, for all $a > 0$,

$$\limsup_{x\to\infty} N[Lw_*(x) - w_*(x)] - rc_1(a,x) < 0.$$

Since (6) holds, it suffices to show that $Lw_*(x) - w_*(x) \leq \mu_F b'(x)$ for all $x \geq 0$. Fix $x \geq 0$ and note that

$$\begin{aligned} Lw_*(x) - w_*(x) &\leq \int_0^\infty \int_x^{x+s} b'(y)\mathrm{d}y\mathrm{d}F(s) && (31)\\ &\leq \int_0^\infty b'(x)s\mathrm{d}F(s) && (32)\\ &= \mu_F b'(x) && (33) \end{aligned}$$

where (31) holds since w_* is absolutely continuous with $w_*' \leq b'$ (because $w_* \in C_*$, from Step 1), and (32) holds since b is concave. Result follows.

Step 4. *If* (7) *holds, then* $\lim_x \sigma_*(x) > 0$. Fix $a > 0$ satisfying (7). Suppose that $\lim_x \sigma_*(x) = 0$ and seek a contradiction. If so then, $\sigma_*(x) \leq a$ for x sufficiently large. Since σ_* solves (5), it suffices to show that

$$\liminf_x N[Lw_*(x) - w_*(x)] - rc_1(a,x) > 0.$$

Let $\hat{b} = \lim_x b'(x)$. Since a satisfies (7), it suffices to show that

$$\liminf_x Lw_*(x) - w_*(x) - \mu_F b'(x) \geq 0$$

Moreover, $\hat{c} = \lim_x c_1(a,x) < \infty$. Let x be large enough so that $\sigma_*(x) \leq a$. Then, for $y \geq x$,

$$\begin{aligned} c_2(\sigma_*(x), y) &= \int_0^{\sigma_*(x)} c_{12}(s,y)\mathrm{d}s && (34)\\ &\leq \sigma_*(x)c_{12}(a,y) && (35)\\ \Rightarrow \quad \int_x^{x+s} c_2(\sigma_*(x),y)\mathrm{d}y &\leq \sigma_*(x)[c_1(a,x+s) - c_1(a,x)] \\ &\leq \sigma_*(x)\hat{c} && (36) \end{aligned}$$

where (34) holds since $c(0,x)=0$, (35) holds since $c_{12}(s,x)$ is increasing in s, and (36) holds since $c_1(a,x)$ is increasing in x and positive. Moreover, from Step 1, w_* is absolutely continuous, and a fixed-point of P. Then, from (26), for $y \geq x$,

$$\begin{aligned} w_*'(y) &= \frac{r[b'(y) - c_2(\sigma_*(y), y)] + N\sigma_*(y)(Lw_*)'(y)}{r + N\sigma_*(y)} \\ &\geq \frac{r[b'(y) - c_2(\sigma_*(y), y)]}{r + N\sigma_*(x)} \qquad (37) \\ &\geq \frac{r[\hat{b} - c_2(\sigma_*(x), y)]}{r + N\sigma_*(x)} \qquad (38) \\ \Rightarrow \quad \int_x^{x+s} w_*'(y)\mathrm{d}y &\geq \frac{sr\hat{b} - \sigma_*(x)\hat{c}}{r + N\sigma_*(x)} \qquad (39) \\ \Rightarrow \quad Lw_*(x) - w_*(x) &\geq \frac{\mu_F r\hat{b} - \sigma_*(x)\hat{c}}{r + N\sigma_*(x)} \end{aligned}$$

where (37) holds since Lw_* is increasing, $b(z) - c(\sigma_*(y), z)$ is increasing in z, and σ_* is decreasing, (38) holds since $b' \geq \hat{b}$ and $c_2(\sigma_*(y), y) \leq c_2(\sigma_*(x), y)$, and (39) follows from (36). Result follows by taking the limit as $x \to \infty$, since $\lim_x \sigma_*(x) = 0$. □

6.5 Proofs for Section 4

Proof of Theorem 2. The proof has several steps.

Step 1. *There exist symmetric MPE $\underline{\sigma}_s$ and $\bar{\sigma}_s$ inducing welfare $\underline{w}_s$ and $\bar{w}_s$ such that, for any symmetric MPE σ_s, $\underline{\sigma}_s \leq \sigma_s \leq \bar{\sigma}_s$ and $\underline{w}_s \leq w_s \leq \bar{w}_s$. Moreover, $\bar{\sigma}_s \leq \sigma_*$.* Let C_s be the set of $w \in C$ such that $b \leq w \leq w_*$, where w_* is maximal welfare (c.f. Theorem 1). Note that, by exerting no effort, any player can achieve payoff $b(x)$ in the game with initial stock x, for all $x \geq 0$. Moreover, her payoff cannot exceed the benchmark $w_*(x)$. Therefore, if σ_s is a MPE inducing welfare w_s, then $w_s \in C_s$.

From (9), if the opponents of Player i play a symmetric Markov profile σ, the Bellman equation for the value function of Player i is

$$rw_s(x) = \max_{a \in [0,1]} r[b(x) - c(a,x)] + [a + (N-1)\sigma_s(x)]\,[Lw_s(x) - w_s(x)] \qquad (40)$$

where L is defined in (14). Fix a symmetric Markov profile σ_s inducing welfare

$w_s \in C_s$. Then σ_s is a MPE if and only if, for all $x \geq 0$, (40) holds and

$$\sigma_s(x) \in \arg\max_{a \in [0,1]} a\left[Lw_s(x) - w_s(x)\right] - rc(a,x). \tag{41}$$

Let $Y = \{(x,l) \in \mathbb{R}^2 : x \geq 0, l \geq Lb(x)\}$ with typical element $y = (x,l)$, and $H : [0,1] \times Y \to \mathbb{R}$ be given by

$$H(a,x,l) = l - b(x) + c(a,x) - (r + aN)c_1(a,x). \tag{42}$$

Claim 2. For all $y \in Y$, $H(\cdot, y)$ is decreasing and has at most one root.

Claim 2 is proved at the end of this proof. Let $q : Y \to [0,1]$ be such that $q(y) = 0$ (resp. $q(y) = 1$) if $H(a,y) < (>)\, 0$ for all $a \in [0,1]$, and let $q(y)$ be the unique root of $H(\cdot, y)$ otherwise.

Claim 3. For $w \in C_s$ and $x \geq 0$, $Qw(x) \equiv q(Lw(x), x)$ is the unique $a' \in [0,1]$ such that

$$a' \in \arg\max_{a \in [0,1]} a\left\{Lw(x) - \bar{T}[a', x, Lw(x)]\right\} - rc(a,x) \tag{43}$$

where $\bar{T}$ is defined in (16).

To prove the claim, note that the derivative with respect to a of the objective on the RHS of (43) is $rH(a', Lw(x), x)/(r + Na')$. Then, Claim 3 follows from Claim 2 and the definition of q. Given $w \in C_s$, and $x \geq 0$, let

$$P_s w(x) = \bar{T}[Qw(x), x, Lw(x)].$$

Claim 4. Equations (40) and (41) hold for all $x \geq 0$ if and only if w_s is a fixed-point of P_s and $\sigma_s = Qw_s$.

To prove the claim, note that, since (2) and (3) are equivalent, for any $x \geq 0$ and $a \in [0,1]$, $w_s(x) = \bar{T}[a, x, Lw_s(x)]$ if and only if

$$rw_s(x) = r[b(x) - c(a,x)] + Na[Lw_s(x) - w_s(x)]. \tag{44}$$

Fix $x \geq 0$. Note that '(40) and (41)' are equivalent to '(44) with $a = \sigma_s^i(x)$

and (41)' which are equivalent to '$w_s = \bar{T}[\sigma_s(x), x, Lw_s(x)]$ and (41)'. Moreover, from Claim 3, '$w_s(x) = \bar{T}[\sigma_s(x), x, Lw_s(x)]$ and (41)' are equivalent to '$w_s(x) = \bar{T}[\sigma_s(x), x, Lw_s(x)]$ and $\sigma_s(x) = Qw_s(x)$', which are equivalent to '$w_s(x) = Pw_s(x)$ and $\sigma_s(x) = Qw_s(x)$'. Since $x \geq 0$ is arbitrary, the claim follows.

Then, it suffices to show that P_s admits lowest and greatest fixed points $\underline{w}_s$ and $\bar{w}_s$ in C_s, and that $Q\underline{w}_s \leq Q\bar{w}_s$. Clearly, $(C_s, \leq)$ is a complete lattice, where '$\leq$' is the point-wise order.[26] Then, from Tarski's fixed-point theorem, P_s admits lowest and greatest fixed points $\underline{w}_s$ and $\bar{w}_s$ in C_s if P_s maps C_s to itself and is order-preserving.

Claim 5. For any $w \in C_s$,

1. for all $x \geq 0$, $Qw(x) \leq \max \arg\max_a \bar{T}[a, x, Lw(x)]$
2. if $v \in C_s$ is such that $v \leq w$, then $Qv \leq Qw$
3. for all $x \geq 0$ and $a \leq Qw(x)$, $\bar{T}[a, x, Lw(x)] \leq \bar{T}[Qw(x), x, Lw(x)]$.

Claim 5 is proved at the end of this proof. To show that P_s maps C_s to itself, fix $w \in C_s$. Then, clearly, $P_s w \in C$. Fix $x \geq 0$. Since $Qw(x) \geq 0$, from Part 2 of Claim 5, $P_s w(x) \geq \bar{T}[0, x, Lw(x)] = b(x)$. Moreover, $P_s w(x) = \bar{T}[Qw(x), x, Lw(x)] \leq Pw(x) \leq Pw_*(x) = w_*(x)$, where the first inequality holds by definition of P, given by (23), and the second holds since P is order-preserving. Since $x \geq 0$ is arbitrary, $P_s w \in C_s$, so P_s maps C_s to itself.

To show that P_s is order-preserving, fix $v \in C_s$ such that $v \leq w$ and $x \geq 0$. From Part 1 of Claim 5, $Qv(x) \leq Qw(x)$. Then, Part 2 of the claim implies that $\bar{T}[Qv(x), x, Lw(x)] \leq P_s w(x)$. Moreover, $P_s v(x) = \bar{T}[Qv(x), x, Lv(x)] \leq \bar{T}[Qv(x), x, Lw(x)]$, where the inequality holds by definition of $\bar{T}$, since $v \leq w$. Hence $P_s w(x) \geq P_s v(x)$ and, since $x \geq 0$ is arbitrary, P_s is order-preserving.

Therefore P_s admits a lowest and greatest fixed points $\underline{w}_s$ and $\bar{w}_s$ in C_s. Note that, since $\underline{w}_s \leq \bar{w}_s$, Part 1 of Claim 5 implies that $Q\underline{w}_s \leq Q\bar{w}_s$. This completes the first step.

Step 2. *If (6) holds, the symmetric MPE is unique.* From Step 1, it suffices to show that $\underline{\sigma}_s = \bar{\sigma}_s$. From Part 1 of Claim 5, it suffices to show that $\underline{w}_s = \bar{w}_s$.

26. For any set $A \subseteq C_s$, the join (resp. meet) of A is $\hat{w}(x) = \sup\{w(x) : w \in A\}$ (resp. $\tilde{w}(x) = \inf\{w(x) : w \in A\}$.).

Claim 6. If $\lim_x \sigma_*(x) = 0$ then $\lim_x w_*(x) - b(x) = 0$.

Claim 7. For all $x \geq 0$

$$\bar{w}_s(x) - \underline{w}_s(x) \leq L\bar{w}_s(x) - L\underline{w}_s(x). \tag{45}$$

Claim 6 is proved at the end of the proof, whereas Claim 7 is proved at the end of the step. Suppose that, for some $\epsilon > 0$ and $x \geq 0$, $\bar{w}_s(x) - \underline{w}_s(x) > \epsilon$, and seek a contradiction. Let

$$x' = \sup\{x \geq 0 : \bar{w}_s(x) - \underline{w}_s(x) \geq \epsilon\}.$$

Since (6) holds for all $a > 0$, from Theorem 1, benchmark effort σ_* is decreasing with $\lim_{x\to\infty} \sigma_*(x) = 0$. Then Claim 6 implies that $\lim_{x\to\infty} w_*(x) - b(x) = 0$. Since $\underline{w}_s, \bar{w}_s \in C_s$, it follows that $\lim_{x\to\infty} \bar{w}_s(x) - \underline{w}_s(x) = 0$. Then $x' < \infty$, and $\bar{w}_s(x) - \underline{w}_s(x) < \epsilon$ for all $x > x'$. But then (45) implies that $\bar{w}_s(x') - \underline{w}_s(x') < \epsilon$, contradicting the definition of x'.[27] Hence $\underline{w}_s(x) = \bar{w}_s(x)$ for all $x \geq 0$.

It remains to prove Claim 7. Fix $x \geq 0$. There are three cases.

Case 1: $\underline{\sigma}_s(x) = 1$. Since $\bar{\sigma}_s \geq \underline{\sigma}_s$, $\bar{\sigma}_s(x) = 1$ also. From Step 1, $w_s(x) = T[1, x, Lw_s(x)]$ for $w_s \in \{\underline{w}_s, \bar{w}_s\}$. Then

$$\bar{w}_s(x) - \underline{w}_s(x) = \frac{N}{r+N}[L\bar{w}_s(x) - L\underline{w}_s(x)]$$

so (45) holds.

Case 2: $\bar{\sigma}_s(x) = 0$. Then, by a similar argument, $\underline{\sigma}_s(x) = 0$, so $w_s(x) = b(x)$ for $w_s \in \{\underline{w}_s, \bar{w}_s\}$. Hence (45) holds.

Case 3: $\underline{\sigma}_s(x) < 1$ *and* $\bar{\sigma}_s(x) > 0$. Note that (41) holds for $(\sigma_s, w_s) \in \{(\underline{\sigma}_s, \underline{w}_s)(\bar{\sigma}_s, \bar{w}_s)\}$.

27. To show this, note that, from (45), $\bar{w}_s(x_0) - \underline{w}_s(x') \leq L\bar{w}_s(x') - L\underline{w}_s(x') \leq \int_0^\infty \bar{w}_s(x'+s) - \underline{w}_s(x'+s)\mathrm{d}F(s) < \epsilon$ where the strict inequality follows from the fact that $F(0) = 0$ and $w_s(x) - \underline{w}_s(x) < \epsilon$ for all $x > x'$.

Then

$$L\bar{w}_s(x) - \bar{w}_s(x) \geq rc_1\left[\bar{\sigma}_s(x), x\right] \tag{46}$$

$$\geq rc_1\left[\underline{\sigma}_s(x), x\right] \tag{47}$$

$$\geq L\underline{w}_s(x) - \underline{w}_s(x) \tag{48}$$

where (46) follows from (41), since $\bar{\sigma}_s(x) > 0$, (47) from the fact that $\bar{\sigma}_s(x) \geq \underline{\sigma}_s(x)$, and (48) from (41), since $\underline{\sigma}_s(x) < 1$. Rearranging yields (45). This proves Claim 7, and completes the proof of Step 2.

Step 3. *σ_s is decreasing.* From Step 2, it suffices to find a symmetric MPE σ_s that is decreasing. From Claim 4, it suffices to exhibit a fixed point $w_s \in C_s$ of P_s such that Qw_s is decreasing.

Let C_s^2 be the set of $w \in C_s$ such that, for any $x_2 \geq x_1 \geq 0$,

$$w(x_2) - w(x_1) \leq b(x_2) - b(x_1).$$

Claim 8. For $w \in C_s^2$, $Lw(x) - b(x)$ is decreasing.

To prove the claim, fix $w \in C_s^2$, $x_2 \geq x_1 \geq 0$. Then, it suffices to show that $Lw(x_2) - b(x_2) \leq Lw(x_1) - b(x_1)$. Note that

$$\begin{aligned}
Lw(x_2) - Lw(x_1) &= \int_0^\infty w(x_2 + s) - w(x_1 + s)\mathrm{d}F(s) \\
&\leq \int_0^\infty b(x_2 + s) - b(x_1 + s)\mathrm{d}F(s) \\
&= \int_0^\infty \int_{x_1}^{x_2} b'(x + s)\mathrm{d}x\mathrm{d}F(s) \\
&\leq \int_0^\infty \int_{x_1}^{x_2} b'(x)\mathrm{d}x\mathrm{d}F(s) \qquad (49) \\
&= b(x_2) - b(x_1)
\end{aligned}$$

where (49) holds since b is concave. The claim follows by rearranging.

Claim 9. For $w \in C_s^2$, Qw is decreasing.

To prove the claim, fix $w \in C_s^2$. Claim 8 implies that $Lw(x) - b(x)$ is decreasing. Note that $c(a, x) - (r + Na)c_1(a, x)$ is decreasing in x, since $c_1, c_{12} \geq 0$. Then

$H[a, x, Lw(x)]$, defined in (42), is decreasing in x. Result follows from Claim 3.

Given Claim 9, it suffices to find a fixed point $w_s \in C_s^2$ of P_s. The argumet used to prove Lemma 2 implies that any chain in C_s^2 admits a join and a meet. Moreover, P_s is order-preserving on C_s^2 since it is order-preserving on C_s. Then, from Lemma 1, it remains to prove that P_s maps C_s^2 to itself. Fix $w \in C_s^2$. From Step 1, $P_s w \in C_s$. Hence, it remains to prove that, for all $x_2 \geq x_1 \geq 0$, $P_s w(x_2) - P_s w(x_1) \leq b(x_2) - b(x_1)$. Fix x_1 and x_2, and let $q_i = Qw(x_i)$ for $i = 1, 2$. Then Claim 9 implies that $q_1 \geq q_2$. Hence

$$P_s w(x_2) = \bar{T}[q_2, x_2, Lw(x_2)] - \bar{T}[q_2, x_1, Lw(x_1)] + \bar{T}[q_2, x_1, Lw(x_1)]$$

$$\Rightarrow \quad P_s w(x_2) - P_s w(x_1) \leq \bar{T}[q_2, x_2, Lw(x_2)] - \bar{T}[q_2, x_1, Lw(x_1)] \tag{50}$$

$$\leq \frac{r[b(x_2) - b(x_1)] + Nq_2[Lw(x_2) - Lw(x_1)]}{r + Nq_2} \tag{51}$$

$$\leq b(x_2) - b(x_1) \tag{52}$$

where (50) follows from Part 2 of Claim 5, since $q_2 \leq q_1$, (51) holds since $c(q, x)$ is increasing in x, and (52) follows from Claim 8. Hence $P_s w \in C_s^2$.

Step 4. *σ_s and w_s are continuous.*

Claim 10. w_s is continuous.

Claim 10 is proved at the end of this proof. It remains to show that σ_s is continuous.

Claim 11. q is continuous.

Claim 12. Lw is continuous for any continuous $w \in C_s$.

Claims 11 and 12 are proved at the end of this proof. From Claim 10, w_s is continuous. Then, since $w_s \in C_s$, Claim 12 implies that Lw_s is continuous. Then, Claim 11 implies that σ_s is continuous. This completes the proof. □

Proof of Claim 2. Note that H has has continuous partial derivatives and

$$H_1(a, x, l) = -(N-1)c_1(a, x) - (r + Nq)c_{11}(a, x).$$

Then, $H(\cdot, x, l)$ is (strictly decreasing) on $[0,1]$ for all $(x,l) \in Y$ (such that $x > 0$). It remains to show that $H(\cdot, y)$ has at most one root for $y \in Y$ such that $y = (0,l)$ for some $l \geq Lb(0)$. Fix l and let $a_0 = \sup\{a \in [0,1] : c_1(a,x) = 0\}$. Then, $H(\cdot, 0, l)$ is strictly decreasing on $[a_0, 1]$. Hence, it suffices to show that $H(a,0,l) > 0$ for all $a < a_0$.

Fix $a < a_0$ and note that $H(a,0,l) = l - b(0)$. Since $l \geq Lb(0)$ and F is not degenerate at 0, it suffices to show that $b(s) > b(0)$ for all $s > 0$. Fix $s > 0$ and note that

$$\begin{aligned} b(s) &\geq b(0) - c(a_0, 0) + c(a_0, s) && (53) \\ &= b(0) + c(a_0, s) && (54) \\ &> b(0) && (55) \end{aligned}$$

where (53) holds since $u(a_0, \cdot)$ is increasing, (54) holds by definition of a_0, since $c(0,0) = 0$, and (55) holds since $c(\cdot, s)$ is strictly increasing for any $s > 0$. □

Proof of Claim 5. To prove the first part, fix $w \in C_s$ and note that $w \geq b$. Then, for any $a \in [0,1]$,

$$\begin{aligned} Lw(x) &\geq Lb(x) \\ &\geq b(x) && (56) \\ &\geq b(x) - c(a,x) && (57) \end{aligned}$$

where (56) holds since b is increasing. Note that $\bar{T}[a, x, Lw(x)]$ is quasi-concave and differentiable with respect to a, and its derivative has the same sign as

$$\begin{aligned} N[Lw(x) - b(x) + c(a,x)] - (r + Na)c_1(a,x) &= H[a, x, Lw(x)] + (N-1)[Lw(x) - b(x) + c(a,x)] \\ &\geq H[a, x, Lw(x)]. \end{aligned}$$

where '$\geq$' follows from (57), and H was defined in (42). Then, result follows by definition of Q and q.

To prove the second part, fix $v \in C_s$ such that $v \leq w$. Then $Lv \leq Lw$. Hence,

by definition of Q, it suffices to show that $q(x,\cdot)$ is increasing for all $x \geq 0$. This follows from the fact that, for all $x \geq 0$ and $a \geq 0$, $H(a,x,\cdot)$ is continuous and increasing.

To prove the third part, fix $x \geq 0$ and $a \leq Qw(x)$. If $Qw(x) = 0$, then since $a \leq Qw(x)$, $a = Qw(x)$ so $\bar{T}[a,x,Lw(x)] = \bar{T}[Qw(x),x,Lw(x)]$. Hence, suppose that $Qw(x) > 0$. Let $l : [0,1] \to \mathbb{R}$ be such that

$$l(q) = T[q, q + (N-1)Qw(x), x, Lw(x)]$$

where T was defined in (15). Note that l is differentiable with

$$l'(q) = r\frac{Lw(x) - b(x) + c(q,x) - [r + q + (N-1)Qw(x)]c_1(q,x)}{[r + q + (N-1)Qw(x)]^2}. \tag{58}$$

Since $c(q,x)$ is convex in q, the nominator of the fraction on the RHS of (58) is decreasing, hence l is quasi-concave. Therefore, l is maximised at any $q \in [0,1]$ such that $l'(q) = 0$. Note that

$$l'[Qw(x)] = \frac{rH[Qw(x),x,Lw(x)]}{[r + NQw(x)]^2}$$

where H was defined in (42). Then, Claim 3 implies that l is maximised at $q = Qw(x)$.

Then, it suffices to show that $l(a) \geq \bar{T}[a,x,Lw(x)]$. Since l is maximised at $b = Qw(x)$ and it is quasi-concave, and $a \leq Qw(x)$, $l'(a) \geq 0$. Since $c_1 \geq 0$, (58) implies that $b(x) - c(a,x) \leq Lw(x)$. Then,

$$\begin{aligned}
\bar{T}[a,x,Lw(x)] &= \frac{r[b(x) - c(a,x)]}{r + Na} + \left[1 - \frac{r}{r + Na}\right] Lw(x) \\
&\leq \frac{r[b(x) - c(a,x)]}{r + a + (N-1)Qw(x)} + \left[1 - \frac{r}{r + a + (N-1)Qw(x)}\right] Lw(x) \\
&= l(a)
\end{aligned}$$

where the inequality holds since $a \leq Qw(x)$. □

Proof of Claim 6. Note that, by construction, $w_* - b$ is decreasing. Hence, for

any $x \geq 0$

$$w_*(x) \leq w_*(0) - b(0) + b(x).$$

Then

$$\begin{aligned} Lw_*(x) &\leq w_*(0) - b(0) + Lb(x) \\ &\leq w_*(0) - b(0) + b(x) + b'(0)\mu_F \end{aligned} \tag{59}$$

where (59) holds since b is concave. Hence

$$\begin{aligned} w_*(x) &= \bar{T}[\sigma_*(x), x, Lw_*(x)] \\ &= \frac{r\{b(x) - c[\sigma_*(x), x]\} + N\sigma_*(x)Lw_*(x)}{r + N\sigma_*(x)} \\ &\leq b(x) + \frac{-rc[\sigma_*(x), x] + N\sigma_*(x)(w_*(0) - b(0) + b'(0)\mu_F)}{r + N\sigma_*(x)} \end{aligned}$$

Therefore, since $\lim_x \sigma_*(x) = 0$,

$$\lim_x w_*(x) - b(x) \leq -\lim_x c[\sigma_*(x), x] \leq 0.$$

Result follows since, by construction, $w_* \geq b$. □

Proof of Claim 10. Given a function $g : [0, \infty) \to \mathbb{R}$, and $x \geq 0$, write

$$g(x^+) = \lim_{y \downarrow x} g(y)$$

$$g(x^-) = \begin{cases} \lim_{y \uparrow x} g(y) & \text{if } x > 0 \\ g(0) & \text{if } x = 0. \end{cases}$$

Since $w_s \in C_s^2$

$$w_s(x_2) - w_s(x_1) \leq b(x_2) - b(x_1) \quad \forall x_2 \geq x_1 \geq 0. \tag{60}$$

Then, since b is continuous, $w_s(x^+)$ and $w_s(x^-)$ exist for all $x \geq 0$, and

$$w_s(x^-) \geq w_s(x) \geq w_s(x^+). \tag{61}$$

Suppose that w_s is discontinuous at some $x_0 \geq 0$ and seek a contradiction. For $x \geq 0$, let

$$l(x) = w_s(x^+) - w_s(x^-).$$

I claim that $l(x_0) < 0$. If $w_s(x_0^+) < w_s(x_0)$ then, from (61), $l(x_0) \leq w_s(x_0^+) - w_s(x_0) < 0$ as desired. If instead $w_s(x_0^+) \geq w_s(x_0)$, then (61) implies that w_s is right-continuous at x_0. Hence, $x_0 > 0$ and w_s is not left-continuous at x_0. Therefore, $w_s(x_0^-) \neq w_s(x_0)$. From, (61), $w_s(x_0) < w_s(x_0^-)$. Then, $l(x_0) = w_s(x_0) - w_s(x_0^-) < 0$. This proves the claim.

Let

$$Y = \{y \geq 0 : l(y) \leq l(x_0)\}.$$

I claim that Y is finite. To prove the claim, fix $N \in \mathbb{N}$, and suppose that $|Y| \geq N$. Then, there is $(y_n)_{n=1}^N \subseteq Y$ such that $y_1 < \cdots < y_N$. Then, by definition of Y,

$$\begin{aligned} w_s(y_N^+) &\leq l(x_0) + w_s(y_N^-) \\ &\leq l(x_0) + b(y_N) - b(y_{N-1}) + w_s(y_{N-1}^+) \qquad (62) \\ &\leq Nl(x_0) + b(y_N) - b(y_1) + w_s(y_1^-) \qquad (63) \end{aligned}$$

where (62) follows from (65), and (64) follows by iterating the argument. Since $b \leq w_s \leq w_*$ and b and w_s are continuous, $b(y_N) \leq w_s(y_N^+)$ and $w_s(y_1^-) \leq w_*(y_1)$. Therefore (64) implies that

$$b(y_1) - w_*(y_1) \leq Nl(x_0). \tag{64}$$

Since b and w_* are continuous and $\lim_x w_*(x) - b(x) = 0$, $b - w_*$ is bounded below. But if Y were infinite, (64) would hold for all $N \in \mathbb{N}$. Since $l(x_0) < 0$, this is impossible. Hence, Y is finite, as desired.

Let $z_0 = \max Y$. Since $z_0 \in Y$, $l(z_0) \leq l(x_0)$. Since we are seeking a contradic-

tion, it suffices to show that $l(z_0) > l(x_0)$. To this end, note that

$$Lw_s(x_2) - Lw_s(x_1) \leq Lb(x_2) - Lb(x_1) \quad \forall x_2 \geq x_1 \geq 0 \tag{65}$$

so that $Lw_s(x^+)$ and $Lw_s(x^-)$ exist for all $x \geq 0$. Then

$$Lw_s(z_0^+) - Lw_s(z_0^-) = \int_0^\infty l(z_0 + s)\mathrm{d}F(s) < l(x_0) \tag{66}$$

where the inequality follows from the fact that $F(0) = 0$ and $l(z_0 + s) < l(x_0)$ for all $s \geq 0$. Consider two cases.

Case 1: $\sigma_s(z_0) = 1$ and σ_s is continuous at z_0. Then

$$\begin{aligned} l(z_0) &= \bar{T}[\sigma_s(z_0), z_0, Lw_s(z_0^+)] - \bar{T}[\sigma_s(z_0), z_0, Lw_s(z_0^+)] \\ &= \frac{N\sigma_s(z_0)}{r + N\sigma_s(z_0)}[Lw_s(z_0^+) - Lw_s(z_0^-)] \end{aligned}$$

Then, since $l(x_0) < 0$, (66) yields $l(z_0) > l(x_0)$, as desired.

Case 2: either $\sigma_s(z_0) < 1$ or σ_s is discontinuous at z_0. Since σ_* is decreasing, $\sigma_s(x) < 1$ for all $x > z_0$, and $\sigma_*(z_0^+)$ and $\sigma_*(z_0^-)$ exist. Fix $x > z_0$. Since $\sigma_s(x) < 1$, from the Bellman equation for optimal effort

$$w_s(x) \geq Lw_s(x) - rc\left[\sigma_s(x), x\right].$$

Since this holds for all $x > z_0$, c is continuous, and $Lw_s(z_0^+)$ and $\sigma_*(z_0^+)$ exist, taking the limit $x \downarrow z_0$,

$$w_s(z_0^+) \geq Lw_s(z_0^+) - rc\left[\sigma_s(z_0^+), z_0\right]. \tag{67}$$

Since $w_s \geq b$ and $l(z_0) \leq l(x_0) < 0$, there is an increasing sequence $(x_n)_{n\in\mathbb{N}}$ converging to z_0 such that $w_s(x_n) > b(x_n)$ for all $n \in \mathbb{N}$. Fix $n \in \mathbb{N}$. Since $w_s(x_n) > b(x_n)$, $\sigma_s(x_n) > 0$. Then, from the Bellman equation for optimal effort,

$$w_s(x_n) \leq Lw_s(x_n) - rc\left[\sigma_s(x_n), x_n\right].$$

Since this holds for all $n \in \mathbb{N}$, c is continuous, and $Lw_s(z_0^-)$ and $\sigma_s(z_0^-)$ exist, taking the limit as $n \to \infty$,

$$w_s(z_0^-) \leq Lw_s(z_0^-) - rc\left[\sigma_s(z_0^-), z_0\right].$$

Then, from (67),

$$\begin{aligned} l(z_0) &\geq Lw_s(z_0^+) - Lw_s(z_0^-) + r\left\{c\left[\sigma_s(z_0^-), z_0\right] - c\left[\sigma_s(z_0^+), z_0\right]\right\} \\ &\geq Lw_s(z_0^+) - Lw_s(z_0^-) \end{aligned} \tag{68}$$

where (68) holds since $c(\cdot, z_0)$ is increasing and σ_s is decreasing. Then, $l(z_0) > l(x_0)$ follows from (66) and (68). □

Proof of Claim 11. Fix y and let $(y_n)_n \subseteq Y$ be such that $\lim_n y_n = y$. It remains to show that $\lim_n Q(y_n) = Q(y)$.

Case 1. $Q(y) \in (0,1)$. In this case, $H[Q(y), y] = 0$. Since $H(\cdot, y)$ is decreasing and has at most one root on $[0,1]$, $H(0,y) > 0$ and $H(1,y) < 0$. Then, since H is continuous, there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $H(0, y_n) > 0$ and $H(1, y_n) < 0$. Fix $n \geq N$. By the intermediate-value theorem, $H(\cdot, y_n)$ has a root in $(0,1)$. Hence, $H[Q(y_n), y_n] = 0$. Since H is continuous, $H(\lim_n Q(y_n), y) = \lim_n H[Q(y_n), y_n] = 0$. Since $H(\cdot, y)$ has at most one root, $\lim_n Q(y_n) = Q(y)$, as desired.

Case 2. $Q(y) \in \{0,1\}$ *and* $H(a,y) \neq 0$ *for all* $a \in [0,1]$. Suppose $Q(y) = 0$ (the case $Q(y) = 1$ is similar). Then, $H(0,y) < 0$. Therefore, since H is continuous, there exists $N \in \mathbb{N}$ such that $H(0, y_n) < 0$ for all $n \geq N$. Fix $n \geq N$. Since H is decreasing, $H(a, y_n) < 0$ for all $a \in [0,1]$. Hence, $Q(y_n) = 0$. Since $n \geq N$ is arbitrary, $\lim_n Q(y_n) = 0 = Q(y)$, as desired.

Case 3. $Q(y) \in \{0,1\}$ *and* $H(Q(y), y) = 0$. Suppose $Q(y) = 0$ (the case $Q(y) = 1$ is similar). Then, since $H(\cdot, y)$ is decreasing and has at most one root, $H(1,y) < 0$. Then, since H is continuous, there exists $N \in \mathbb{N}$ such that $H(1, y_n) < 0$ for all $n \geq N$. Fix $n \geq N$. Since $H(1, y_n) < 0$, either $H[Q(y_n), y_n] = 0$ or $Q(y_n) = 0$. Consider a subsequence $(y'_n)_n$ such that $\lim_n Q(y'_n)$ exists. It suffices to show

that $\lim_n Q(y'_n) = 0$. Suppose that $\lim_n Q(y'_n) > 0$, and seek a contradiction. If $\lim_n Q(y'_n) > 0$, there exists $N' \in \mathbb{N}$ such that $H[Q(y'_n), y'_n] = 0$ for all $n \geq N'$. But then, since H is continuous, $H[\lim_n Q(y'_n), y] = 0$, so $\lim_n Q(y'_n) = Q(y) = 0$, a contradiction. Hence $\lim_n Q(y'_n) = 0 = Q(y)$, as desired. □

Proof of Claim 12. Fix a continuous $w \in C_s$. Fix $x \geq 0$ and $\epsilon > 0$. We seek $\delta > 0$ such that $|Lw(x) - Lw(y)| \leq \epsilon$ for all $y \geq 0$ such that $|x - y| \leq \delta$. Note that, since $\lim_x w_*(x) - b(x) = 0$, there exists $\hat{x} \geq 0$ such that $w_*(y) - b(y) \leq \epsilon/3$ for all $y \geq \hat{x}$. Then, since $b \leq w \leq w_*$ (as $w \in C_s$) and w_* and b are increasing, $|w(y) - w(z)| \leq w_*(y) - b(z) \leq b(y) - b(z) + \epsilon/3$ for all $y \geq z \geq \hat{x}$. Note that b is Lipschitz. Hence, there exists $\delta_1 > 0$ such that, for all $y, z \geq 0$ such that $|y - z| \leq \delta_1$, $|b(y) - b(z)| \leq \epsilon/3$. Then, $|w(y) - w(z)| \leq 2\epsilon/3$ for all $y, z \geq \hat{x}$ such that $|y - z| \leq \delta_1$.

If $x \geq \hat{x} + \delta_1$, let $\delta = \delta_1$. Then, for all $y \geq 0$ such that $|x - y| \leq \delta$, $y \geq \hat{x}$, so

$$|Lw(x) - Lw(y)| \leq \int_0^\infty |w(x+s) - w(y+s)| \mathrm{d}F(s) \leq \frac{2\epsilon}{3}.$$

Suppose instead that $x < \hat{x} + \delta_1$. Since w_s is continuous, it is uniformly continuous on $[0, \hat{x} + \delta_1]$. Then, there exists $\delta_2 > 0$ such that $|w(y) - w(z)| < \epsilon/3$ for all $y, z \in [0, \hat{x} + \delta_1]$ such that $|y - z| \leq \delta_2$. Let $\delta = \min\{\delta_1, \delta_2\}$, and fix $y \geq 0$ such that $|y - x| \leq \delta$. Then $x + s, y + s \leq \hat{x}$ for all $s \leq \hat{x} - x + \delta$ Moreover, $x + s, y + s \geq \hat{x}$ for all $s \geq \hat{x} - x - \delta$. Hence

$$\begin{aligned}
|Lw(x) - Lw(y)| &\leq \int_0^\infty |w(x+s) - w(y+s)| \mathrm{d}F(s) \\
&\leq \int_0^{\hat{x}-x+\delta} |w(x+s) - w(y+s)| \mathrm{d}F(s) + \int_{\hat{x}-x-\delta}^\infty |w(x+s) - w(y+s)| \mathrm{d}F(s) \\
&\leq \frac{\epsilon}{3} F(\hat{x} - x + \delta) + \frac{2\epsilon}{3}[1 - F(\hat{x} - x - \delta)] \leq \epsilon.
\end{aligned}$$

Hence, Lw is continuous, as desired. □

Proof of Proposition 1. The second part was proved at the end of Section 4.1. Hence, it suffices to show that w_s is increasing if $c(a, x) = c(a)$. Suppose that $c(a, x) = c(a)$. From Step 2 of the proof of Theorem 2, it suffices to find a symmetric

MPE σ_s such that w_s is increasing. From Claim 4, it suffices to exhibit a fixed point $w_s \in C_s$ of P_s such that w_s is increasing.

Let C_s^3 be the subset of absolutely continuous $w \in C_s$ such that $0 \leq w' \leq b'$. Then, it suffices to show that P_s has a fixed point in C_s^3. Then, Lemma 2 implies that any chain in C_s^3 has a join and a meet. Since P_s is order-preserving on C_s and $C_s^3 \subseteq C_s$, Lemma 1 implies that it suffices to show that P_s maps C_s^3 to itself.

Claim 13. If $w \in C_s^2$ is absolutely continuous, then Qw_s and $P_s w$ are absolutely continuous and, for all $x > 0$ such that $(Qw)'(x) \in (0,1)$,

$$(Qw)'(x) = \frac{(Lw)'(x) - b'(x) + c_2(q,x) - (r+Nq)c_{12}(q,x)}{(N-1)c_1(q,x) + (r+Nq)c_{11}(q,x)} \tag{69}$$

$$(P_s w)'(x) = \frac{r\delta(q) + Nq}{r+Nq}(Lw)'(x) + \frac{r[1-\delta(b)]}{r+Nq}[b'(x) - c_2(q,x)] - r\delta(q)c_{12}(q,x) \tag{70}$$

where $q = Qw(x)$, and

$$\delta(q) = \left[1 + \frac{r+Nq}{N-1}\frac{c_{11}(q,x)}{c_1(q,x)}\right]^{-1}. \tag{71}$$

Claim 13 is proved at the end of this proof. Fix $w \in C_s^3$. From Claim 13, $P_s w$ is absolutely continuous. Moreover, from Step 2 of the proof of Theorem 2, $P_s w \in C_s^2$. Then, $(P_s w)' \leq b'$. Hence, it suffices to show that $(P_s w)'(x) \geq 0$ for all $x > 0$.

Fix $x > 0$. If $Qw(y) = 0$ in a neighbourhood of x, then $(P_s w)'(x) = b'(x) \geq 0$. If $Qw(y) = 1$ in a neighbourhood of x, then

$$(P_s w)'(x) = \frac{rb'(x) + N(Lw)'(x)}{r+N} \geq 0$$

where the inequality holds since $(Lw)'(x) \geq 0$, as w is increasing. Finally, if $Qw(x) \in (0,1)$ in a neighbourhood of x, (70) holds so that, setting $q = Qw(x)$,

$$(P_s w)'(x) = \frac{r\delta(q) + Nq}{r+Nq}(Lw)'(x) + \frac{r[1-\delta(b)]}{r+Nq}b'(x) \geq 0.$$

Result follows. □

Proof of Claim 13. Fix an absolutely continuous $w \in C_s^2$ and let $I = \{x \geq 0 :$

$Qw(x) \in (0,1)\}$. To prove that Qw is absolutely continuous and that (69) holds at any $x \in I$, note that Qw is decreasing from Step 2 of the proof of Theorem 2. Then, it suffices to show that Qw is absolutely continuous on I. Since $w \in C_s^2$ is absolutely continuous, $w' \leq b'$. Then, from the argument used to derive (25), Lw is absolutely continuous with $(Lw)' \leq b'$. Fix $x \in I$. By definition of q, $Qw(x)$ is the unique root of $H(\cdot, x, Lw(x))$ in $[0,1]$. Since H is continuously differentiable in a neighbourhood of x, result follows from the implicit function theorem.

Since $P_s w = \bar{T}[Qw(x), x, Lw(x)]$ where $\bar{T}$ is continuously differentiable, $P_s w$ is absolutely continuous since Qw and Lw are. Then, it remains to show that, for any $x > 0$ such that $Qw(x) \in (0,1)$, (70) holds. Fix x and let $q = Qw(x)$. Since $q \in (0,1)$, Claim 3 implies that $H[q, x, Lw(x)] = 0$, where H is defined in (42). Rearranging gives

$$Lw(x) - b(x) + c(q,x) = (r + Nq)c_1(q,x)$$

Then

$$\begin{aligned}\bar{T}_1[q, x, Lw(x)] &= \frac{-rc_1(q,x)(r+Nq) + rN[Lw(x) - b(x) + c(q,x)]}{(r+Nq)^2} \\ &= \frac{r(N-1)c_1(q,x)}{r+Nq}.\end{aligned}$$

Then, from (69),

$$\bar{T}_1[q, x, Lw(x)](Qw)'(x) = \frac{r\delta(q)}{r+Nq}\left[(Lw)'(x) - b'(x) + c_2(q,x) - (r+Nq)c_{12}(q,x)\right].$$

Finally, note that

$$(P_s w)'(x) = \bar{T}_1[q, x, Lw(x)](Qw)'(x) + \frac{r[b'(x) - c_2(q,x)] + Nq(Lw)'(x)}{r+Nq}$$

Result follows by substituting the expression for the first term, and rearranging. □

Chapter 3

Incentives for Collective Innovation

Payoff-maximising Equilibria and Concealment

Abstract

I carry on the analysis of the model presented in Chapter 2. I characterise the payoff-maximising equilibrium for the two-agent binary-effort case, when strategies may condition on time and the past trajectory of the stock, and show that it is asymmetric. I then allow agents inducing increments in the stock to conceal them from other agents for a period of time, forgoing their benefits. In this case, in the symmetric equilibrium with full disclosure studied in Chapter 2, agents would conceal innovations that are not sufficiently beneficial. I construct an equilibrium in which innovations that do not exceed a common cutoff are concealed, and show that it is more efficient than the one with full disclosure. In the context of an R&D partnership in which improvements in a shared technology raise the incentive to shift resources towards private activities, concealing moderate improvements may raise private profits and collective progress.

1 Introduction

In this chapter, I continue the analysis of the dynamic game introduced in Chapter 2. In this model, identical agents exert hidden effort to induce randomly-sized increments in the stock of a public good, and their opportunity cost of effort is increasing in the stock. The game captures instances of collective innovation in which progress decreases the incentive to innovate further.

I present two main results. First, in the two-agent case with binary effort, the 'anonymous' public-perfect equilibrium that maximises payoffs is asymmetric and non-Markov. Only one agent exerts effort after the stock reaches a certain cutoff and, ex-ante, agents are equally likely to be the ones to do so. Second, if agents may conceal the innovations they obtain (forgoing higher payoffs) for a period of time, they would conceal moderate ones until, after further improvements, they become significant. Doing so is beneficial to everyone.

Payoff-maximising equilibrium. I characterise the subgame-perfect equilibrium maximising aggregate payoffs of the two-player game with binary effort, assuming that strategies may only condition on time and the past trajectory of the stock. In this sense, the equilibrium is anonymous. Since effort is hidden, the equilibrium is inefficient. Nevertheless, aggregate payoffs are higher than in the unique symmetric Markov equilibrium after every history. In fact, they are maximised (within the aforementioned class of equilibria) after every history.

Aggregate equilibrium effort is pinned down by the current stock. In particular, time and the past trajectory of the stock do not provide useful incentives. However, the equilibrium is not Markov, as the stock's past trajectory is a useful correlation device. Only one agent exerts effort after the stock exceeds a cutoff and, before this occurs, each agent has a one-in-two chance of being the one to do so. In particular, the equilibrium is (ex-post) asymmetric. Equilibrium effort ceases after the stock exceeds the single-agent cutoff and, when this occurs, aggregate payoffs drop unless the last innovation is significant.

Game with concealment. Since the opportunity cost of effort increases with the stock, innovations may be detrimental in the unique symmetric Markov equilibrium (analysed in Chapter 2), as well as in the payoff-maximising anonymous equilibrium discussed above.

I alter the baseline model by allowing any agent generating an increment in the stock to forgo the larger payoff flow, and conceal the innovation from other agents. At any later time, possibly after having obtained and concealed further innovations, she may disclose any portion of the total increment, increasing her payoffs. I show that the original Markov equilibrium continues to exist if and

only if equilibrium welfare is sufficiently increasing. In particular, delaying the disclosure of moderate innovations may be beneficial as it preserves the opponents' incentives to exert effort. I conjecture that, whether or not the Markov equilibrium continues to exist, the game with concealment admits a symmetric perfect-Bayesian equilibrium in which innovations are concealed until they exceed a common cutoff. After this occurs, all further innovations are immediately disclosed, and agents play the original equilibrium (which continues to exist if the stock is sufficiently high). I expect the new equilibrium to be more efficient than the old one, as the detrimental effect of moderate innovations on the incentives to exert effort is reduced.

1.1 Literature review (continued from Chapter 2)

The analysis of the payoff-maximising equilibria when effort is unobservable contributes to the literature on dynamic public-good games. Previous work focused on specific procedures for provision (e.g. Admati and Perry 1991), asymptotic efficiency (e.g. Marx and Matthews 2000) or on symmetric equilibria (e.g. Battaglini, Nunnari, and Palfrey 2014). Instead, I compute the anonymous equilibrium of the two-player game with binary effort inducing highest aggregate payoffs. This equilibrium is asymmetric, as are the asymptotically-efficient equilibria of Marx and Matthews (2000). However, unlike in Marx and Matthews (2000), asymptotic efficiency cannot be achieved in equilibrium. We will see that this because production is stochastic.

The game with concealment is related to the large literature on information disclosure in the context of innovation. Bimpikis, Ehsani, and Mostagir (forthcoming) study the design of information spillovers within contests. Halac, Kartik, and Liu (2016) consider the joint design of information spillovers and reward schemes. Akcigit and Liu (2015) instead analyse the strategic concealment of 'dead-ends' among competing innovators. Campbell, Ederer, and Spinnewijn (2014) consider a setting in which agents exert hidden effort, obtain increments after random times, and may conceal them in order to preserve their partners' effort incentives. However, their model does not account for the possibility that concealing findings from partners may result in redundancy.

The rest of the paper is organised as follows. In Section 2, I characterise the anonymous subgame-perfect equilibrium maximising aggregate payoffs in the two-player game with binary effort. In Section 3, I analyse the game with concealment. The conclusion is in Section 4.

2 Payoff-Maximising Equilibrium

In this section, I characterise the payoff-maximising equilibrium when strategies may condition on time and the past trajectory of the stock, and show that it takes a simple form. The equilibrium is more efficient than the unique symmetric Markov perfect equilibrium (MPE) described in Section 4 of Chapter 2. Throughout this section, I restrict attention to the two-agent case of the specific example introduced in Chapter 2 (and described below).

As attention is restricted to pure strategies that condition on public histories alone, there is no need to keep track of agents' beliefs about their opponents' private information (i.e, their past effort choices). For this reason, I use Public-Perfect Equilibria (PPE) as the solution concept. Moreover, these strategies ensure that the path of play is unique, even though time is continuous. That is because each agent can affect the public history only by obtaining an innovation, and only finitely many can be obtained in any bounded interval of time.

Assumption 1. Suppose that $N = 2$, that $F(s) = 1 - e^{-\lambda s}$ for some $\lambda > 0$, that $b(x) = x$ and $c(a, x) = a\alpha x$ where $\alpha \in (0, 1]$.

The analysis in Section 3 of Chapter 2 shows that, when cost is linear, effort in the single-agent problem ceases when the stock reaches the cutoff $\hat{y}$ given by

$$\hat{y} = \frac{1}{r\lambda\alpha}. \tag{1}$$

The following result shows that the encouragement effect does not arise in any public-perfect equilibrium. This is because agents have no control over the size of the innovations, and can only alter their frequency.

Lemma 1 (Encouragement effect)**.** *Suppose that Assumption 1 holds. Then no PPE features the encouragement effect. Namely effort is exerted until x_t exceeds $\hat{y}$, given in (1), in the single-agent problem, and no effort is exerted if $x_t > \hat{y}$ in any PPE.*

Lemma 1 is proved in the Appendix. I first show that, in any PPE σ, effort cannot be sustained above some cutoff $y \geq 0$. Then, if the stock x_t is arbitrarily close to, but below y, the next innovation is arbitrarily likely to push the stock above y. In this case, no innovation will follow. Hence, the incentive to exert effort at x_t in σ is essentially no stronger than in the single-agent setting. Therefore, $y = \hat{y}$.

I restrict attention to strategies that condition on time and on the past trajectory of the stock. In particular, they neglect the identity of the past innovators. Hence, label PPE in these strategies *Anonymous-Perfect Equilibria (APE)*. APE allow for a richer set of behaviours than MPE, as the latter compel effort to depend on the current stock alone. In particular, APE may in principle generate stronger incentives to exert effort by having agents play a less efficient equilibrium if the stock does not increase for a sufficiently long period of time, or if it has increased too slowly in the past. We will see that this device is not useful: in the most efficient APE, aggregate effort is pinned down by x_t. The next result characterises the most efficient APE of the game. It is proved in the Appendix.

Theorem 1 (Anonymous PPE)**.** *Suppose that Assumption 1 holds and that*

$$(1+r)[\alpha(1+r)+1]\log\left(1+\frac{1}{\alpha(1+r)}\right) \geq 1. \tag{2}$$

Then there exists an APE σ_a maximising welfare among APE after every history. Aggregate equilibrium effort e_a is unique and is a function of the stock x_t. Agents exert simultaneous effort initially, then only one is active, and eventually stops. Continuation payoffs are symmetric when both agents are active, and welfare w_a may drop when effort ceases.

The existence of an equilibrium that maximises payoffs after any history is by no means guaranteed. In fact, existence often fails even if the initial stock x_0 is fixed

and attention is restricted to histories on the equilibrium path. This is because one can provide strong incentives to act efficiently by having agents play an inefficient continuation equilibrium after histories that are more likely to be reached after inefficient play. Yet, this notion of efficiency is necessary to be able to maximise aggregate payoffs in a time-consistent manner. If agents may co-ordinate on a different equilibrium in each period, then the notion of efficiency based on expected discounted payoffs at the beginning of the game is inappropriate.

Examples of equilibria that maximise payoffs after every history are rare in the literature on dynamic games. One reason is that their existence is ruled out if the trajectory of the state is not monotone.[1] An example is given in Keller, Rady, and Cripps (2005), in which the state is essentially monotone. They characterise the MPE of the two-player game in pure strategies involving finitely many switches between the safe and the risky arm, and show that they can be ranked according to use of the risky arm uniformly across states. Moreover, equilibria involving more experimentation at all states are more efficient. This implies the existence of a most efficient equilibrium.

The argument in the proof of Theorem 1 is different. I do not characterise the full set of APE, as it is rich.[2] Instead, I show that the highest achievable welfare in APE, as a function of the initial state x_0, ($w(x_0)$, say) lies below the solution of a particular Bellman equation. I then show that this solution corresponds to the welfare function of a particular APE σ_a. The main step in the construction of the equation is to note that whenever Agent i exerts effort in some APE σ at some time t, her time-t continuation payoff v_i satisfies

$$\int_0^\infty \bar{v}_i(x_t + s)\mathrm{d}F(s) - v_i \geq r\alpha x_t$$

where $\bar{v}_i(x_t + s)$ is her continuation payoff after an innovation of size s occurs at

1. See the discussion in Keller and Rady (2010).
2. In particular, equilibria cannot be ranked according to effort or welfare after every history.

time t. If agents exert effort simultaneously then this holds for $i = 1, 2$, so that

$$\begin{aligned}\frac{v_1 + v_2}{2} - r\alpha x_t &\leq \int_0^\infty \frac{1}{2}[\bar{v}_1(x_t + s) + \bar{v}_2(x_t + s)]\mathrm{d}F(s) \\ &\leq \int_0^\infty w(x_t + s)\mathrm{d}F(s). \end{aligned} \tag{3}$$

If there exist a sequence of APE σ's in which simultaneous effort is exerted, and inducing welfare that approximate $w(x_t)$, then (3) implies that

$$w(x_t) - r\alpha x_t \leq \int_0^\infty w(x_t + s)\mathrm{d}F(s).$$

If no such sequence exists, then since $N = 2$, $w(x_t)$ must be approximated by APE in which only one agent exerts effort at stock size x_t. The fact that the cost of effort is linear implies that this agent must exert maximal effort. This induces another bound on $w(x_t)$. Combining the bounds produces the Bellman equation.

Even though APE allow behaviour to depend on time and on the whole past trajectory of the stock, aggregate effort and welfare in the most efficient APE σ_a are pinned down by x_t. In particular, time and the past trajectory of x_t do not provide useful direct incentives for effort. Yet, σ_a is not a MPE. That is because the scope for co-ordination is larger in APE than in MPE, and this matters for welfare. In particular, in σ_a, only one agent exerts (maximal) effort when the stock lies within an interval $[\hat{z}, \hat{y}]$, and her identity is not a function of the current stock. Rather, when the stock is smaller than $\hat{z}$, each agent faces a one-in-two chance of having to exert effort in the interval $[\hat{z}, \hat{y}]$ once the stock enters it.[3] This device generates symmetric continuation payoffs at stocks smaller than $\hat{z}$, and this is crucial to maximise effort, and therefore welfare. In other words, the payoff-maximising APE is also the most equitable APE from an ex-ante perspective.[4]

Figure 1 depicts aggregate effort and welfare in the payoff-maximising APE σ_a.

3. Markov strategies do not allow this because the probability of the stock increasing to some $x \in [\hat{z}, \hat{y}]$ varies with their current size. Moreover, a fair lottery can only be induced if F is atomless.

4. The same is true in Keller, Rady, and Cripps (2005).

Aggregate effort satisfies

$$e_a(x) = \begin{cases} 1 & \text{if } x \in [\hat{z}, 1/(r\alpha\lambda)] \\ 0 & \text{if } x > 1/(r\alpha\lambda). \end{cases}$$

Welfare satisfies

$$w_a(x) = \begin{cases} \frac{1}{2r\lambda}\left[\alpha - \frac{1}{r+1}\right] e^{\frac{r\lambda x - 1/\alpha}{1+r}} + \left(x + \frac{1}{r\lambda}\right)\left(1 - \frac{\alpha}{2}\right). & \text{if } \hat{z} \leq x \leq 1/(r\alpha\lambda) \\ x & \text{if } x > 1/(r\alpha\lambda). \end{cases} \tag{4}$$

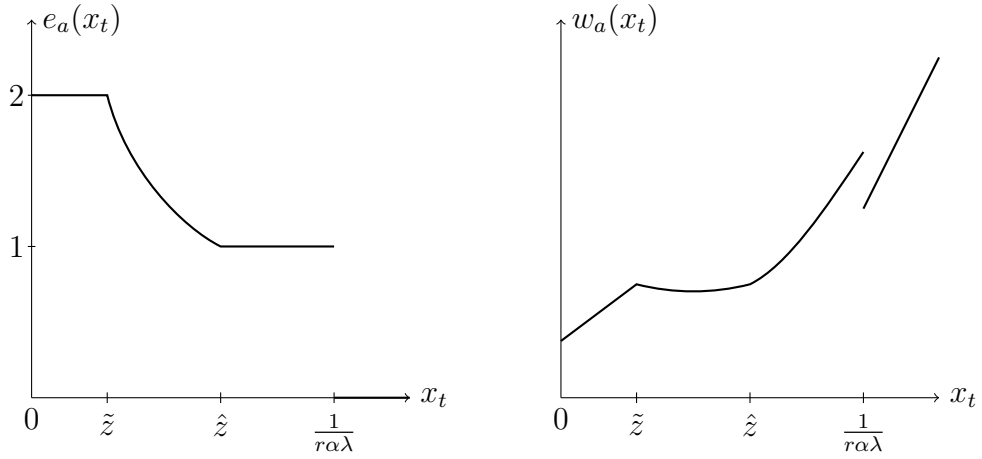


Figure 1: Aggregate effort (left) and welfare (right) in the payoff-maximising anonymous-perfect equilibrium of the example, as functions of the stock x_t.

Non-monotonicity. Welfare w_a is not monotone, but for a different reason than the one breaking monotonicity in the symmetric MPE σ_s. Welfare in σ_s is not monotone if the incentive to exert some effort is arbitrarily large when the stock is small and bounded otherwise (i.e. $c_1(0,0) = 0$ and $c_1(0,x) > 0$ for some $x > 0$), whereas w_a is not monotone essentially because the cost of effort is linear. In particular, if payoffs are given by $b(x) - c(a,x) = x - a$, then welfare is monotone in σ_s, but it is not monotone in σ_a. Note that, as shown in Figure 1, σ_a is similar to σ_s below $\hat{z}$, and it equals the solution to the single-agent problem above $\hat{z}$.[5] The latter

5. This equilibrium is qualitatively similar to the Markov equilibrium described in Section 6 of Keller and Rady (2010). Welfare is monotone in their equilibrium because, in games of experimentation, welfare and the incentive to exert effort are increasing in the state. In contrast, in classical public good games as well as in this model, the incentive to exert effort is decreasing in the state.

is more efficient than any symmetric equilibrium in a neighbourhood of $\hat{y}$ if the cost of effort is not too convex (in particular, if it is linear). Welfare drops at $\hat{y}$ because the continuation payoffs of the agent exerting effort in $[\hat{z}, \hat{y}]$ are continuous in a neighbourhood of $\hat{y}$ and, since she ceases to exert effort abruptly, the continuation payoffs of her opponent drop.

3 Concealment

In this section, I study the impact of allowing agents to conceal the innovations that they induce, after observing their size. Among other things, I argue how this may improve welfare over σ_s.

The baseline model is modified as follows. As before, x_t denotes the stock of the public good and, in any interval of time $[t, t + \mathrm{d}t)$, Agent i exerts effort a_t^i and obtains payoffs $[b(x_t) - c(a_t^i, x_t)\mathrm{d}t$. For any $t \geq 0$ and $i \leq N$, let $k_t^i \geq x_t$ denote Agent i's *private stock* of the good. Initially, x_t and $k_t^1, \ldots, k_t^N$ equal x_0. In any interval of time $[t, t + \mathrm{d}t)$, Agent i produces an increment in k_t^i (instead of x_t) with probability $a_t^i \mathrm{d}t$. The size of innovations has distribution F, as before.

Moreover, any Agent i such that $k_t^i > x_t$ may *disclose* a part $d_t^i \in (x_t, k_t^i]$ of her private stock. If one or more agents disclose at time t, x_t increases to the value $\max_i d_t^i$ of the largest disclosure, along with any $k_t^j < \max_i d_t^i$. Assume that, in each period, agents choose effort, then increments in private stocks are realised, then agents choose whether and how much to disclose. To ensure that the path of play is unique, assume also that given some $\epsilon > 0$ and for any $t > 0$, if the last disclosure by any agent prior to time t occurred at time $t' < t$, any Agent i can disclose at time t if and only if either she obtained an innovation since time t' (i.e. $k_t^i > k_{t''}^i$ for some $t' < t'' < t$), or $t' = t + n\epsilon$, for some integer $n > 0$.

Notice that agents may disclose any amount d_t^i of their private stock k_t^i. If this were not possible, e.g. if agents could only disclose the full stock k_t^i (or nothing), an agent may have an incentive to disclose k_t^i as evidence to her opponents that she is not concealing more substantial innovations. Doing so would raise the opponents' incentives to exert effort, and benefit the agent. Hence, unraveling would occur in

equilibrium, leading all agents to fully, immediately disclose all innovations.

The social-welfare benchmark (c.f. Theorem 1 in Chapter 2) is unaffected by the introduction of concealment. This is because, in the non-strategic setting, the effects of delayed disclosure on incentives can be ignored. Moreover, delayed disclosure hinders the growth of the payoff flow, since $b(x) - c(a, x)$ is increasing in x, and may lead to duplicate innovations. I state the result for completeness, and omit the proof as it is straightforward.

Proposition 1. *The social-welfare benchmark in the game with concealment is identical to the baseline one. In particular, it is efficient to disclose all innovations immediately.*

The rest of this section focuses on Perfect Bayesian Equilibria (PBE). I restrict attention to PBE in which agents' behaviour is a function of the publicly-available information (i.e. the public history) and the current value of their private stock. That is, agents disregard the past trajectory of their private stock, as well as their past effort choices. This is a natural assumption, as the information that is neglected is payoff-irrelevant. Moreover, PBE of this type involve relatively simple beliefs, that merely keep track of the current levels of the opponents' private stocks.

Note that the standard notion of PBE, defined for discrete-time games, applies to this setting. Indeed, since effort is hidden, a public history only records (the current time and) the sequence of past disclosures, involving the times at which they occurred, their values, and the identities of the agents performing them. Given the restriction of agents' behaviour, it suffices to consider information sets recording the public history as well as the current value of the agents' private stock. View agents' actions as specifying their behaviour until either their private stock increases (due to an innovation, or a disclosure) or they disclose. Then, since the arrival of innovations is a Poisson process with bounded rate, and given the restriction on agents' disclosures, agents reach finitely many different information sets within any bounded interval of time. This implies that discrete-time PBE are well-defined in this game.

In this setting, equilibrium *welfare* is the average of agents' expected discounted payoffs at $t = 0$, given that it is common knowledge that $k_0^i = x_0$ for all $i \leq N$.

3.1 Symmetric MPE with Full Disclosure

This section derives conditions under which the symmetric MPE σ_s derived in Section 3 of Chapter 2, continues to be an equilibrium in the game with concealment.

Suppose that agents exert effort according to σ_s, and each one discloses fully their innovations, as soon as they arise. Refer to this strategy profile simply as 'σ_s' for the remainder of this section. If σ_s is played and Agent i obtains an innovation raising her private stock to $k_t^i > x_t$, disclosing k_t^i raises the flow of benefits from $b(x_t)$ to $b(k_t^i)$, but decreases the opponents' efforts from $\sigma_s(x_t)$ to $\sigma_s(k_t^i)$. Hence, Agent i may have an incentive to conceal her innovations. The next result characterises when σ_s is an equilibrium in the game with concealment. It is proved in the Appendix.

Lemma 2. *Given $x_0 \geq 0$, the following are equivalent:*

1. *The MPE σ_s (with full, immediate disclosure by all agents) is a PBE.*
2. *For all $k \geq x \geq x_0$, if σ_s is played, $x_t = x$, and $k_t^i = k$, delaying the (full) disclosure of k_t^i by a small time period dt is unprofitable for Agent i.*
3. *For all $k \geq x \geq x_0$, $\Delta(k, x) \leq 0$ where*

$$\Delta(k,x) = \max_{a\in[0,1]} r[b(x) - c(a,x)] + a\int_0^\infty w_s(k+s) - w_s(k)dF(s) + (N-1)\sigma_s(x)\int_{k-x}^\infty w_s(x+s) - w_s(k)dF(s) - rw_s(k). \quad (5)$$

The value $\Delta(k, x)$ is the net benefit from delaying the disclosure of $k_t^i = k$ by $\mathrm{d}t$ when $x_t = x$ (divided by $\mathrm{d}t$). The first three terms are the continuation payoffs from delaying the disclosure, whereas $rw_s(k)$ is the payoff from disclosing immediately. The payoff from delaying the disclosure, and exerting optimal effort a within the period $[t, t + \mathrm{d}t)$, is an expression similar to the Bellman equation for welfare w_s in the MPE.[6] Namely, it is the sum of the payoff flow $b(x) - c(a, x)$, and the next expected future benefits. These benefits are the second and third terms. The second term is the benefit from the event that Agent i obtains an innovation raising her private stock to $k + s$, and then discloses. In this case, her continuation

6. That is: $rw_s(x) = \max_{a\in[0,1]} r[b(x) - c(a,x)] + [a + (N-1)\sigma_s(x)]\int_0^\infty w_s(x+s) - w_s(x)\mathrm{d}F(s)$.

payoffs are $w_s(k+s)$. The third term arises from the event that an opponent obtains an innovation raising her stock to some value $x+s>k$. In this case, Agent i's continuation payoffs are $w_s(x+s)$. Note that, if the increment s is small so that $x+s<k$, the continuation payoffs are $w_s(k)$ since the Agent i's disclosure exceeds that of her opponent. Moreover, the event that more than one agent obtain an innovation within $[t,t+\mathrm{d}t)$ can be neglected, because the probability that it occurs is small.

The next result shows that, if the cost of effort does not increase with the stock, then concealing innovations is never profitable.

Proposition 2. *If the cost of effort does not increase with the stock, i.e. $c(a,x)=c(a)$, then the symmetric MPE σ_s with full, immediate disclosure by all agents is a PBE.*

Proof. Fix $x_0\geq 0$. From Lemma 2, it remains to show that $\Delta(k,x)\leq 0$ for all $k\geq x\geq x_0$. Note that $\Delta(x_0,x_0)=0$ since w_s satisfies the Bellman equation in footnote 6. Hence, it suffices to show that $\Delta(k,x_0)$ is decreasing in k on $[x_0,\infty)$. From Proposition 1 in Chapter 2, w_s is increasing. Then, the third and fourth terms in (5) are decreasing in k. Hence, it suffices to show that the integral in the second term is decreasing in k. Since σ_s is continuous and decreasing, the first-order condition for optimal effort implies that there are $0\leq\tilde{y}\leq\hat{y}\leq\infty$ such that $\sigma_s(x)=0$ for $x>\hat{y}$,

$$rc'\left[\sigma_s(x)\right]=\int_0^\infty w_s(x+s)-w_s(x)\mathrm{d}F(s) \tag{6}$$

for $x\in[\tilde{y},\hat{y}]$, and $\sigma_s(x)=1$ for $x<\tilde{y}$.

For $x>\hat{y}$, since $\sigma_s(x)=0$, $w_s(x)=b(x)$, so result follows since b is concave. For $x\in[\tilde{y},\hat{y}]$, result follows from (6), since σ_s is decreasing, and c' is increasing. Finally, for $x<\tilde{y}$,

$$\int_0^\infty w_s(x+s)-w_s(x)\mathrm{d}F(s)=\frac{r}{r+N}\int_0^\infty w_s(x+s)-b(x)+c(1)\mathrm{d}F(s).^{7}$$

7. This is because $w_s(x)=\{r[b(x)-c(a)]+Na\int_0^\infty w_s(x+s)\mathrm{d}F(s)\}/(r+Na)$, where $a=\sigma_s(x)$. In turn, this follows from equation (2) in Chapter 2.

The argument in Step 3 of the proof of Theorem 2 in Chapter 2 shows that $w_s(y) - b(y)$ is decreasing. Moreover, for all $s \geq 0$, $b(x+s) - b(x)$ is decreasing in x since b is concave. Setting $y = x + s$ implies that $w_s(x+s) - b(x)$ is decreasing in x. Result follows. □

To analyse the case in which cost increases with the stock, I restrict attention to linear payoffs and increments with finite variance. That is, I make the following assumption. Let

$$\mu_{F^2} = \int_0^\infty s^2 \mathrm{d}F(s)$$

Assumption 2. $\mu_{F^2} < \infty$ and, for some $\alpha \in (0,1]$, $c(a,x) = a\alpha x$, and $b(x) = x$.

The next proposition characterises σ_s and w_s in this particular case. It is proved in the Appendix. Let $g_F : [0,\infty) \to [1,\infty)$ be given by

$$g_F(z) = 1 + \mathbb{E}\left[\max\left\{m \in \mathbb{N} : \textstyle\sum_{i=1}^m S_i < z\right\}\right]$$

where $S_1, S_2, \ldots$ are i.i.d. random variables with distribution F. Note that g_F depends on F, but does not depend on any other parameter.

Proposition 3. *Suppose that Assumption 2 holds. Then effort in σ_s ceases as the stock reaches*

$$\hat{y} = \frac{\mu_F}{r\alpha}$$

and is maximal while the stock lies below $\tilde{y} < \hat{y}$, given by

$$\frac{1}{\tilde{y}} \int_{\tilde{y}}^{\hat{y}} g_F\left(\hat{y} - x\right) dx = N - 1. \tag{7}$$

Moreover

$$\sigma_s(x) = \frac{r}{N-1} \frac{1}{x} \int_x^{\hat{y}} g_F\left(\hat{y} - x\right) dx \quad \forall x \in [\tilde{y}, \hat{y}] \tag{8}$$

$$w_s(x) = \begin{cases} x + r\alpha \int_x^{\hat{y}} g_F\left(\hat{y} - z\right) dz & \text{if } x \in [\tilde{y}, \hat{y}) \\ x & \text{if } x \geq \hat{y}. \end{cases} \tag{9}$$

Let

$$\tilde{x}_F = \frac{\mu_{F^2}}{2\mu_F^2}.$$

The next result follows from Proposition 3, and is relevant for the analysis of the game with concealment. It is proved in the Appendix.

Corollary 1. *The cutoff $\tilde{y}$ is decreasing in N with $\lim_N \tilde{y} = 0$. On $[\tilde{y}, \infty)$, w_s is convex and independent of N, and generally has a unique minimum $\tilde{x}$ such that*

$$\lim_{r\alpha \to 0} \lim_{N \to \infty} \tilde{x} = \tilde{x}_F.^{8} \tag{10}$$

As noted in Corollary 1, for all $x_0 < \tilde{x}_F$, w_0 is non-monotone on $[x_0, \infty)$ if α is sufficiently small (so that, for fixed r, $r\alpha$ is small) and, given α, N is sufficiently large. In this case, agents have an incentive to conceal innovations that would cause welfare w_s to drop. That is because any agent obtaining an innovation can mimick the scenario in which that innovation had not occurred. Moreover, by doing so her payoffs are exactly those that she would earn in this scenario. In particular, agents can avoid any drop in welfare. Hence w_s is not an equilibrium.

In contrast, for any $x_0 > \tilde{x}_F$, σ_s *is* an equilibrium, provided the cost of effort increases sufficiently slowly (i.e. $\alpha > 0$ is sufficiently small). This is intuitive given that, from Proposition 2, σ_s is an equilibrium if the cost does not increase. These results are summarised in the next proposition. It is proved in the Appendix.

Proposition 4. *Suppose that Assumption 2 holds, and that F has unbounded support. If $x_0 < \tilde{x}_F$ then the profile σ_s with full, immediate disclosure by all agents is not an equilibrium in the game with concealment if α is sufficiently small and, given α, the number of agents N is sufficiently large. If $x_0 > \tilde{x}_F$, then σ_s is a PBE if α is sufficiently small and, given α, N is sufficiently large.*

Interestingly, there exists an initial stock x_0 such that w_s is increasing on $[x_0, \infty)$, so that each innovation is 'beneficial' when σ_s is played, and yet agents may have an incentive to conceal. In particular, this is the case if w_s is minimised at x_0, and $\sigma_s(x_0) \in (0, 1)$. Indeed, $\Delta(x_0, x_0) = 0$ and, by the envelope theorem, $\Delta(k, x_0)$ is

8. If F is discrete, w_s may be minimised over an interval for some 'knife-edge' values of $r\alpha$.

continuously differentiable in k on $(x_0, x_0+\epsilon)$ for some $\epsilon > 0$ with

$$\Delta_1(k, x_0) = a(k) \int_0^\infty w_s(k+s) - w_s(k) \mathrm{d}F(s) - \{(N-1)\sigma_s(x)[1-F(k-x)]+r\} w_s'(k)$$

where $a(k)$ is the maximand in (5). Since $c(a, x)$ is increasing in x, $a(k) \geq \sigma_s(k)$. Moreover, since σ_s is continuous, for $\epsilon > 0$ small enough, $\sigma_s(k) \in (0, 1)$, so that (6) implies that the integral equals the marginal cost $r\alpha$. Since w_s is minimised at x_0, $w_s'(x_0) = 0$. It follows that $\lim_{k \downarrow x_0} \Delta_1(k, x_0) = \sigma_s(x_0) r\alpha > 0$. Hence, $\Delta(x_0 + \epsilon, x_0) > 0$ for $\epsilon > 0$ small enough. In this case, as long as the stock has not increased, an agent who obtains a sufficiently small innovation would rather conceal it.

3.2 Symmetric PBE with Concealment

In this section, I conjecture that, if the cost of effort increases sufficiently slowly with the stock, and there are sufficiently many agents, the game with concealment admits a relatively simple PBE, whether or not w_s is increasing and σ_s is a PBE. Moreover, welfare in this PBE exceeds w_s. I maintain Assumption 2 throughout the section.

As shown in Proposition 4, given any $x_0 < \tilde{x}_F$, for α sufficiently small and N sufficiently large, agents have an incentive to conceal innovations that are sufficiently small when σ_s is played. I conjecture that there exists a symmetric PBE σ_d in which, initially, agents conceal innovations as long as their private stocks k_t^i lie below a common, time-varying cutoff $d(t)$. Then, after the first disclosure occurs, agents revert to playing σ_s. The cutoff $d(t)$ exceeds $\tilde{x}_F$ so that, from sition 4, σ_s is an equilibrium for any stock levels exceeding it. Moreover σ_d is more efficient than σ_s in the sense that the sum of agents' expected discounted payoffs at the beginning of the game (when it is common knowledge that $k_t^i = x_0$ for all $i \leq N$), is higher than $w_s(x_0)$. The conjecture is summarised below.

Conjecture 1 (Concealment PBE)**.** *Suppose that Assumption 2 holds. Given $x_0 < \tilde{x}_F$, if $\alpha > 0$ is sufficiently small and, given α, the number of agents is sufficiently large, there exists a symmetric PBE σ_d in which innovations are concealed unless*

they exceed a common, time-varying cutoff $d(t) > \tilde{x}_F$ *and, after the first disclosure,* σ_s *is played. Moreover, initial welfare in* σ_d *is greater than in* σ_s*.*

The equilibrium σ_d is symmetric in the sense that, on path, in any time period, agents with the same private stocks exert the same effort and, at any point in time, hold the same anonymous beliefs about the prive stocks of their opponents. Namely, for all $i \leq N$ and $t \geq 0$, on path,

$$a_t^i = \begin{cases} a_t\left(k_t^i\right) & \text{if } x_t = x_0 \\ \sigma_s(x_t) & \text{otherwise} \end{cases}$$

for some $a : [x_0, \infty) \to [0, 1]$, since $x_t = x_0$ if and only if no agent disclosed prior to or at time t. Moreover, at any time $t \geq 0$ prior to the first disclosure, from the perspective of Agent i, the current private stocks $(k_t^j)_{j \neq i}$ of her opponents are i.i.d. with some CDF β_t over $[x_0, d(t)]$ which does not depend on k_t^i or i. The initial belief β_0 is degenerate with value x_0 and the law of motion of the CDF β_t is given by

$$\dot{\beta}_t(k) = \beta_t(k) \int_{x_0}^{d(t)} a(l)[1 - F(d(t) - l)]\mathrm{d}\beta_t(l) - \int_{x_0}^{k} a(l)[1 - F(k - l)]\mathrm{d}\beta_t(l).^{9}$$

As a consequence, letting v_t^i be the time-t on-path continuation payoffs of Agent i given her private information,

$$v_t^i = \begin{cases} v_t\left(k_t^i\right) & \text{if } x_t = x_0 \\ w_s(x_t) & \text{otherwise} \end{cases}$$

for some $v : [x_0, \infty) \to \mathbb{R}$.

Figure 2 depicts the equilibrium continuation payoffs v_t prior to disclosure, as a function of the private stock k_t^i. I expect v_t to be increasing in k_t^i, as agents

9. This is derived by noting that, if time were discrete with length $\Delta > 0$,

$$\beta_{t+\Delta}(k) = \frac{\int_{x_0}^{k} 1 - \left[1 - e^{-\Delta a(l)}\right][1 - F(k - l)]\mathrm{d}\beta_t(l)}{\int_{x_0}^{d} 1 - \left[1 - e^{-\Delta a(l)}\right][1 - F(d(t) - l)]\mathrm{d}\beta_t(l)}$$

Then, the expression for $\dot{\beta}_t(k)$ is given by $\lim_{\Delta \to 0} \mathrm{d}\beta_{t+\Delta}(k)/\mathrm{d}\Delta$.

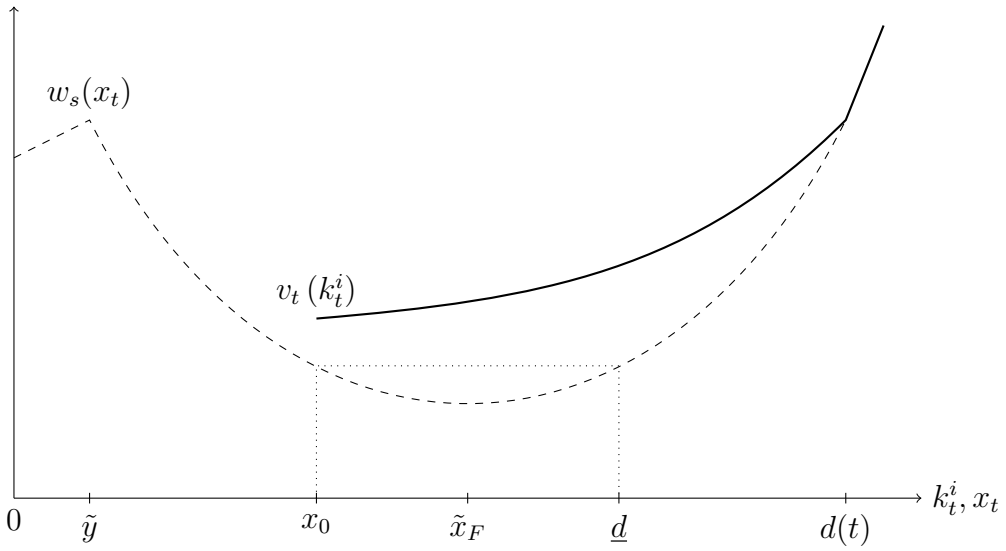


Figure 2: Payoffs $v_t(k_t^i)$ in the PBE σ_d (solid line), and in the MPE $\sigma_s(x_t)$ (dashed line), as functions of the private stock k_t^i and the public stock x_t, respectively. The stock levels have the same scale.

with larger private stocks can achieve the payoffs of agents with smaller ones, by replicating their behaviour. Moreover, the disclosure cutoff $d(t)$ exceeds the value $\underline{d}$, which is the unique $x > \tilde{x}_F$ such that $w_s(x_0) = w_s(x)$. This ensures that, for α sufficiently small, σ_s is an equilibrium for all $x_0 \geq \underline{d}$ (c.f. Proposition 4). Moreover d is such that $v_t(k) \geq w_s(k)$ for all $k \in [x_0, d(t)]$, with equality if $k = d(t)$. In particular, $v_0(x_0) \geq w_s(x_0)$. That is, welfare in σ_d exceeds welfare in σ_s. Also, prior to the first disclosure, an Agent i such that $k_t^i = d(t)$ is indifferent between disclosing k_t^i and not disclosing at time t. Finally, as v_t is increasing in k_t^i, it follows that the first disclosure is beneficial to all agents. Since w_s is increasing above $\underline{d}$, it follows that *all* disclosures are beneficial to all agents.[10]

The equilibrium σ_d involves punishments for any Agent i disclosing a value d_t^i at time t that is strictly below the cutoff $d(t)$. One natural 'punishment' for a disclosure $d_t^i \in (x_0, \tilde{x}_F)$ would be a continuation equilibrium of the same type as σ_d, with the initial stock equal to d_t^i. Figure 2 suggest that this continuation equilibrium would involve payoffs $\hat{v}_t(k)$ strictly lower than $v_t(k)$ for all pairs (k, t). Despite this, it may not be an effective punishment for sufficiently large t. Indeed,

10. Campbell, Ederer, and Spinnewijn (2014) analyse a setting in which agents exert hidden effort, obtain increments after random times, and may conceal them in order to preserve their partners' effort incentives. However, their model does not account for the possibility that concealing findings from partners may result in redundancy.

if $\hat{v}_t(k) < v_t(k)$ for all k then the cutoff $\hat{d}(t)$ of the continuation equilibrium would satisfy $\hat{d}(t) < d(t)$. Hence, immediately after Agent i discloses, an Agent j with private stock $k_t^j \in [\hat{d}(t), d(t)]$ would disclose it, raising Agent i's payoffs. The prospect of this larger disclosure by some Agent j may prompt the deviation of Agent i, despite the risk of lower continuation payoffs $\hat{v}_t(k_t^i)$.[11]

If, at some time t prior to the first disclosure, Agent i discloses $d_t^i < d(t)$, the continuation equilibrium is such that the other agents exert no effort, and disclose nothing, for the remainder of the game. This equilibrium can be supported with the following beliefs. After the unexpected disclosure at time t, any Agent $j \neq i$ believes that k_t^i is very large, and that Agent i will disclose it at time $t + \epsilon$. Note that, since Agent i disclosed at time t, Agent i (and none of her opponents) can disclose before time $t + \epsilon$, unless she obtained a private increment before then. Hence, if indeed k_t^i were large, and Agent i did not obtain a private increment within $(t, t + \epsilon)$, Agent i would act as each of her opponents predicts. However, if $k_t^i < d(t)$, then Agent i would not meet the expectations of her opponents at time $t+\epsilon$. Then, suppose that her opponents would expect a large disclosure from Agent i at time $t + 2\epsilon$. This reasoning can be repeated, and justifies the sudden drop in effort of the opponents of Agent i at time t.

4 Conclusion

In this chapter, I characterised the payoff-maximising anonymous-perfect equilibrium of a specific example of the game introduced in Chapter 2 and showed that it is asymmetric, despite the fact that agents are identical. I also analysed a modified version of the baseline game in which agents may conceal the innovations that they obtain, forgoing their payoffs. The rationale behind concealment is that it preserves partners' incentives to exert effort. I showed that, if the opportunity cost of effort is increasing in the stock, concealing innovations may be beneficial. As a consequence, the symmetric Markov equilibrium studied in Chapter 2 may no

11. If it were possible to force agents to disclose then, in a game where Agent i is forced to disclose k_t^i if and only if $k_t^i > \underline{d}$, the above punishment would be effective. That is because, the disclosure cutoff $d(t)$ would be fixed at $\underline{d}$, and not drop after the disclosure d_t^i.

longer exist. Instead, I conjectured the existence of a perfect Bayesian equilibrium in which innovations are concealed until they reach a cutoff, and disclosed thereafter. Moreover, would be equilibrium is more efficient than the original one, as the detrimental effects of moderate innovations are avoided. In the context of an R&D parternship, these results imply that concealing moderate innovations may raise private profits and collective progress, as it preserves the partners' incentives to R&D.

References

Admati, A., and M. Perry. 1991. "Joint projects without commitment." *The Review of Economic Studies* 58 (2): 259–276.

Akcigit, Ufuk, and Qingmin Liu. 2015. "The role of information in innovation and competition." *Journal of the European Economic Association* 14 (4): 828–870.

Battaglini, M, S Nunnari, and T R Palfrey. 2014. "Dynamic free riding with irreversible investments." *American Economic Review* 104 (9): 2858–71.

Bimpikis, Kostas, Shayan Ehsani, and Mohamed Mostagir. Forthcoming. "Designing Dynamic Contests." *Mathematics of Operations Research.*

Campbell, Arthur, Florian Ederer, and Johannes Spinnewijn. 2014. "Delay and deadlines: Freeriding and information revelation in partnerships." *American Economic Journal: Microeconomics* 6 (2): 163–204.

Halac, Marina, Navin Kartik, and Qingmin Liu. 2016. "Optimal contracts for experimentation." *The Review of Economic Studies* 83 (3): 1040–1091.

Keller, G., and S. Rady. 2010. "Strategic experimentation with Poisson bandits." *Theoretical Economics* 5 (2): 275–311.

Keller, G., S. Rady, and M. Cripps. 2005. "Strategic experimentation with exponential bandits." *Econometrica* 73 (1): 39–68.

Marx, L., and S. Matthews. 2000. "Dynamic voluntary contribution to a public project." *The Review of Economic Studies* 67 (2): 327–358.

5 Appendix

The following piece of notation is used throughout the appendix. Given two functions f and g defined over the same domain X, write '$f \geq g$' for '$f(x) \geq g(x)$ for all $x \in X$'.

A function $w : [0, \infty) \to \mathbb{R}$ is *k-Lipschitz* for some $k > 0$ if, for all $x_1, x_2 \geq 0$,

$$|w(x_2) - w(x_1)| \leq k|x_2 - x_1|.$$

Let C be the set of measurable functions $w : [0, \infty) \to \mathbb{R}$ such that, for some $a, b, c \in \mathbb{R}$ with $a \leq c$ and $b \geq 0$, $w(x) \in [a, bx + c]$ for all $x \geq 0$. For $w \in C$, let

$$Lw(x) \equiv \int_0^\infty w(x + s)\mathrm{d}F(s).^{12} \tag{11}$$

and let $\bar{T} : [0, 1]^2 \times [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be given by

$$\bar{T}(a, x, l) = \frac{r[b(x) - c(a, x)] + Nal}{r + Na}. \tag{12}$$

5.1 Proofs for Section 2

Given a Markov profile σ and $x \geq 0$, let $v^i_\sigma(x)$ (resp. $w_\sigma(x)$) be the payoffs of Agent i (resp. welfare) in the game with initial stock x, given that σ is played. For $x \geq 0$, let H_x be the set of public histories leading to stock level $x \geq 0$. Given a profile σ of strategies that are functions of public histories, and $h \in H_x$, let $w_\sigma(h)$ be welfare in the game starting after history h, when σ is played.

12. Note that the integral exists since w is bounded below by a, and it takes a finite value since it lower than $b(x + \mu_F)$, and $\mu_F < \infty$.

Proof of Lemma 1. It suffices to show that no effort can be sustained above $\hat{y}$ in any PPE. Fix a PPE σ and let x' be the highest stock level x such that effort is exerted above x by some agent, after some history. Clearly $x' < \infty$. Then, there exists $x \leq x'$ arbitrarily close to x' such that, after some history, some Agent i exerts some positive amount of effort before the state increases. Then, her incentive to exert effort is no larger than if her opponent exerted no effort as long as the state does not increase, and would then exert full effort as long as the state lies below x'. Clearly, for $x < x'$ sufficiently close to x', the incentive in this case is only marginally stronger than when playing alone. Then, if $x' > \hat{y}$, Agent i would rather exert no effort, contradicting the definition of x. Hence $x' \leq \hat{y}$, so that $x \leq \hat{y}$, as desired. □

Proof of Theorem 1. The proof has two steps.

Step 1. *Construction of* σ_a. Let $\hat{\sigma}$ be the symmetric Markov profile such that

$$\hat{\sigma}(x) = \begin{cases} 1/2 & \text{if } x \leq \hat{y} \\ 0 & \text{otherwise.} \end{cases}$$

where $\hat{y}$ is given in Lemma 1. Let

$$\hat{z} = \inf\{x \geq 0 : r\alpha(x) \geq Lw_{\hat{\sigma}}(x) - w_{\hat{\sigma}}(x)\}.$$

Let σ_a be such that, after any history $h \in H_x$ where $x \geq \hat{z}$, some Agent 1 exerts full effort until the state exceeds $\hat{y}$, and then stops, whereas Agent 2 exerts no effort.

Claim 1. For any $x_0 \geq \hat{z}$ and any $h \in H_{x_0}$, σ_a is an equilibrium after h.

Claim 1 is proved at the end of this proof. If $\hat{z} = 0$, then σ_a is fully-defined.

If $\hat{z} > 0$, then I construct σ_a such that, the expected continuation payoffs of each agent when the stock exceeds $\hat{z}$ are given by $w_{\hat{\sigma}}$. This is done as follows.

Let $\bar{\sigma}$ be the Markov profile such that $\bar{\sigma}^1(x) = 0$ for all $x \geq 0$, and

$$\bar{\sigma}^2(x) = \begin{cases} 1 & \text{if } x \leq \hat{y} \\ 0 & \text{otherwise.} \end{cases}$$

For $x \geq 0$, let $\bar{v}_i(x)$ be the payoffs of Agent i in the game with initial stock x when $\bar{\sigma}$ is played. Note that $(\bar{v}_1 + \bar{v}_2)/2 = \hat{w}$ since $\bar{\sigma}^1 + \bar{\sigma}^2 = 2\hat{\sigma}$. Moreover, $\bar{v}_1 \geq \bar{v}_2$ where the inequality is strict for all $x \leq \hat{y}$. Then, for all $x < \hat{y}$, there is a unique $s(x) > 0$ such that

$$L\hat{w}(x) = \int_0^{s(x)} \bar{v}_1(x+s)\mathrm{d}F(s) + \int_{s(x)}^{\infty} \bar{v}_2(x+s)\mathrm{d}F(s).$$

Moreover, it is clear that the function $x \mapsto s(x)$, defined on $[0, \infty)$, is measurable. Let σ_a be such that, if an increment occurs when the stock takes value $x \leq \hat{z}$, the continuation equilibrium is given by $\bar{\sigma}$ if the size of the innovation exceeds $s(x)$, and it is given by $\bar{\sigma}$ with the roles of Agent 1 and Agent 2 reversed otherwise. This ensures that, upon an innovation raising the stock x above $\hat{z}$, the expected continuation payoffs of both agents are given by $L\hat{w}(x)$.

It remains to specify σ_a at histories leading to stock levels below $\hat{z}$. I construct σ_a such that, below $\hat{z}$, effort is symmetric and is a function of the current stock level alone. This is done by adapting the argument used to prove Step 3 of Theorem 2 in Chapter 2. Namely, note that $x \leq \hat{w} \leq w_*$, and that $\hat{w}' \leq 1$.[13] Hence, $\hat{w}$ can be extended to some $w \in C_s^2$. Modify the map P_s as follows. For $w \in C_s^2$, let

$$\hat{P}_s w = \begin{cases} P_s w(x) & \text{if } x \leq \hat{z} \\ w(x) & \text{if } x > \hat{z}. \end{cases}$$

It is easy to show that $\hat{P}_s$ has the same properties of P_s, and that the argument in Step 3 of Theorem 2 in Chapter 2 continues to hold. This implies the existence of a symmetric and decreasing Markov profile $\hat{\sigma}_s$ such that, if the payoffs of each agent after the stock exceeds $\hat{z}$ are given by w_{σ_a}, then $\hat{\sigma}_s$ is an equilibrium for any $x_0 \leq \hat{z}$. Moreover, by definition of $\hat{z}$, $\hat{\sigma}_s(\hat{z}) = 1/2$. Let σ_a be such that, after any history leading to stock $x \leq \hat{z}$, both agents exert effort $\hat{\sigma}_s(x)$. This completes the

13. The latter can be shown by noting that (from (2) in Chapter 2),

$$\hat{w}(x) = \frac{rx(1-\alpha/2) + L\hat{w}(x)}{r+1}$$

for $x \in [\hat{z}, \hat{y}]$, and $\hat{w}(x) = x$ for $x \geq \hat{y}$.

construction of σ_a.

Moreover, it is clear that σ_a is an equilibrium, and that it satisfies the properties of the statement of Theorem 1. Then, the next step completes the proof.

Step 2. *The APE σ_a induces higher welfare than any other APE, after any history.* Let Σ be the set of all APE. For all $x \geq 0$, let H_x be the set of public histories leading to the stock level x. Remember that, for any $\sigma \in \Sigma$ and $h \in H$, $w_\sigma(h)$ is welfare in σ after history h. Let

$$\psi_\sigma(x) = \sup_{h \in H_x} w_\sigma(h).$$

$$\psi(x) = \sup_{\sigma \in \Sigma} \psi_\sigma(x)$$

Let $\hat{T} : [0, \infty)^2 \to \mathbb{R}$ be given by

$$\hat{T}(x, l) = \max \left\{ \bar{T}(1/2, x, l), \min \left\{ l - r\alpha x, \max_a \bar{T}(a, x, l) \right\} \right\}.$$

Claim 2. For any $\sigma \in \Sigma$ and $x \leq \hat{y}$, $\psi_\sigma(x) \leq \hat{T}[x, L\psi(x)]$.

To prove the claim, fix $\sigma \in \Sigma$ and $x \geq 0$. By definition of $\psi_\sigma(x)$, $\psi_\sigma(x) = \lim_n w_\sigma(h^n)$ for some sequence $(h_n)_n \subset H_x$. Let $\mathrm{d}t > 0$, and consider the game after some h_n, assuming that no innovation occurs before time $\mathrm{d}t$. Note that, either both agents exert effort at some time $t < \mathrm{d}t$, or aggregate effort does not exceed 1 before time $\mathrm{d}t$. This implies that we may consider the two cases below.

Case 1. There exists a subsequence $(h'_n) \subseteq (h_n)$ such that, for all n and $dt > 0$, both agents exert effort at some time $t < dt$ in the game after h'_n, if no innovation occurs before time dt. In this case, the argument in Section 3 implies that, for all n,

$$w_\sigma(h'_n) - r\alpha(x) \leq L\psi(x).$$

Taking the limit as n diverges and rearranging implies that

$$\psi_\sigma(x) \leq L\psi(x) - r\alpha(x). \tag{13}$$

Case 2. There exists a subsequence $(h'_n) \subseteq (h_n)$ such that, for all n and $dt > 0$,

aggregate effort does not exceed 1 prior to time dt in the game after h'_n, if no innovation occurs before time dt. I claim that, in this case

$$\psi_\sigma(x) \leq \bar{T}[1/2, x, L\psi(x)]. \tag{14}$$

Fix n and $\mathrm{d}t > 0$, and consider the game after h'_n. Let $a_{n,\mathrm{d}t} \in [0, 1/2]$ be average effort per agent over the period $[0, \mathrm{d}t)$, let $l_{n,\mathrm{d}t} \in \mathbb{R}$ be expected continuation welfare after the first innovation, conditional on it occurring before time $\mathrm{d}t$, and let $h_{n,\mathrm{d}t} \in H_x$ be the history reached at time $\mathrm{d}t$ if no innovation occurs. Then

$$\begin{aligned} w_\sigma(h'_n) &\approx r[b(x) - a_{n,\mathrm{d}t}\alpha(x)]\mathrm{d}t + (1 - r\mathrm{d}t)\left[2a_{n,\mathrm{d}t}\mathrm{d}t l_{n,\mathrm{d}t} + (1 - 2a_{n,\mathrm{d}t}\mathrm{d}t)w_\sigma(h_{n,\mathrm{d}t})\right] \\ &\approx r[b(x) - a_{n,\mathrm{d}t}\alpha(x)]\mathrm{d}t + 2a_{n,\mathrm{d}t}\mathrm{d}t l_{n,\mathrm{d}t} + (1 - r\mathrm{d}t)w_\sigma(h_{n,\mathrm{d}t}) - 2a_{n,\mathrm{d}t}\mathrm{d}t w_\sigma(h_{n,\mathrm{d}t}) \end{aligned}$$

where the last line follows by neglecting terms in $\mathrm{d}t^2$. Rearranging gives

$$\begin{aligned} \frac{w_\sigma(h'_n) - w_\sigma(h_{n,\mathrm{d}t})}{\mathrm{d}t(r + 2a_{n,\mathrm{d}t})} + w_\sigma(h_{n,\mathrm{d}t}) &= \bar{T}(a_{n,\mathrm{d}t}, x, l_{n,\mathrm{d}t}) \\ &\leq \bar{T}[a_{n,\mathrm{d}t}, x, L\psi(x)] \end{aligned}$$

where the inequality holds since $l_{n,dt} \leq L\psi(x)$. I claim that, since $a_{n,\mathrm{d}t} \leq 1/2$,

$$\bar{T}[a_{n,\mathrm{d}t}, x, L\psi(x)] \leq \bar{T}[1/2, x, L\psi(x)]$$

Note that $\psi_\sigma \geq w_{\sigma_a}$, since $\sigma_a \in \Sigma$. Then $L\psi(x) \geq Lw_{\sigma_a}(x)$. If $x \leq \hat{z}$, then $\sigma_a(h'_n) = \hat{\sigma}_s(x) \geq \hat{\sigma}_s(\hat{z}) = 1/2$, since $\hat{\sigma}$ is decreasing on $[0, \hat{z}]$. Since $a_{n,\mathrm{d}t} \leq 1/2$, the argument used to prove Part 3 of Claim 5 of the proof of Theorem 2 in Chapter 2 can be adapted to show that

$$\bar{T}[a_{n,\mathrm{d}t}, x, Lw_{\sigma_a}(x)] \leq \bar{T}[1/2, x, Lw_{\sigma_a}(x)].$$

Then, since $\bar{T}(a, x, l)$ is increasing in l, result follows since $L\psi(x) \geq Lw_{\sigma_a}(x)$. If $x \geq \hat{z}$, then result follows by the same argument, since $\bar{T}[a, x, Lw_{\sigma_a}(x)]$ is increasing

in a.[14] Therefore,

$$\frac{w_\sigma(h'_n) - w_\sigma(h_{n,\mathrm{d}t})}{\mathrm{d}t(r + 2a_{n,\mathrm{d}t})} + w_\sigma(h_{n,\mathrm{d}t}) \leq \bar{T}[1/2, x, L\psi(x)]. \tag{15}$$

Since this holds for all n, taking the limit of the supremum,

$$\limsup_n \frac{\psi_\sigma(x) - w_\sigma(h_{n,\mathrm{d}t})}{\mathrm{d}t(r + 2a_{n,\mathrm{d}t})} + w_\sigma(h_{n,\mathrm{d}t}) \leq \bar{T}[1/2, x, L\psi_\sigma(x)]. \tag{16}$$

Clearly, $\psi_\sigma(x) \geq w_\sigma(h) \geq x$ for all $x \geq 0$ and $h \in H_x$. Then, since the above inequality holds for all $\mathrm{d}t > 0$,

$$\psi_\sigma(x) = \lim_{\mathrm{d}t \to 0} \lim_n w_\sigma(h_{n,\mathrm{d}t}).$$

Then, by definition of $\psi_\sigma(x)$, $\psi_\sigma(x) \geq w_\sigma(h_{n,\mathrm{d}t})$, so that the limit of the LHS of (16) is at least $\psi_\sigma(x)$. Hence (14) holds. Therefore, either (13) or (14) hold. Then, by definition of $\hat{T}$, it suffices to show that $\psi_\sigma(x) \leq \max_a \bar{T}[a, x, L\psi(x)]$. This holds since, clearly, given that $x_0 = x$ and continuation payoffs after the first innovation are given by ψ, welfare is maximised by exerting some amount of symmetric, constant effort before the first innovation.

Let $\hat{y} = 1/(r\alpha\lambda)$ and, for all $x \geq 0$, let $w_*(x)$ be the maximal welfare achievable in the game with initial stock x (c.f. Theorem 1 in Chapter 2). Let C_a be the set of $\psi \in C$ such that $x \leq \psi(x) \leq w_*(x)$ for all $x \geq 0$, and $\psi(x) = x$ for all $x \geq \hat{y}$.

Claim 3. For any $\sigma \in \Sigma$, $\psi_\sigma \in C_a$.

To prove the claim, fix $\sigma \in \Sigma$. Then $\psi_\sigma \leq w_*$ by definition of w_*. Moreover, for any $x \geq 0$, agents exerting no effort in the game with initial stock x earn payoffs at least x, hence $\psi_\sigma(x) \geq x$. Finally, from Lemma 1, $\psi_\sigma(x) = x$ for all $x > \hat{y}$. Hence, $\psi_\sigma \in C_a$. This proves the claim.

Note that, since $\bar{T}(a, x, l)$ is increasing in l, then so is $\hat{T}$. Moreover

Claim 4. For all $x \geq 0$, $\hat{T}(x, l)$ is 1-Lipschitz in l.

14. This can be seen by substituting (4) for w_{σ_a}.

For $\psi \in C_a$ and $x \geq 0$ let

$$P_a\psi(x) = \begin{cases} x & \text{if } x > \hat{y} \\ \hat{T}[x, L\psi(x)] & \text{if } x \leq \hat{y}. \end{cases}$$

Claim 5. P_a maps C_a to itself.

To prove the claim, fix $\psi \in C_a$. Since $x \leq \hat{y}$,

$$\begin{aligned} & x + \mu_F - x \geq r\alpha x \\ \Leftrightarrow \quad & x + \mu_F - r\alpha x \geq x. \end{aligned}$$

Moreover, $L\psi(x) \geq x + \mu_F$ since $\psi(y) \geq y$ for all $y \geq 0$. Then, $L\psi(x) - r\alpha x \geq x$. Moreover, $\max_a \bar{T}[a, x, L\psi(x)] \geq x$. Then, $\min\{L\psi(x) - r\alpha x, \max_a \bar{T}[a, x, L\psi(x)]\} \geq x$. Then, $P_a\psi(x) \geq x$ follows by definition of P_a. This proves the claim.

Note that C_a is a complete lattice with respect to the point-wise order. Moreover, since $\hat{T}(x, l)$ is increasing in l for all $x \geq 0$, and the operator L is order-preserving, then so is P_a. Then, by Tarski's fixed-point thereom, P_a admits greatest and lowest fixed-points $\bar{\psi}$ and $\underline{\psi}$ in C_a.

Claim 6. P_a has a unique fixed-point ψ_a in C_a.

To prove the claim, it suffices to show that $\bar{\psi} = \underline{\psi}$. Suppose that for some $x \geq 0$, $\bar{\psi}(x) > \underline{\psi}(x)$, and seek a contradiction. Let

$$x_0 = \sup\{y \geq 0 : \bar{\psi}(y) - \underline{\psi}(y) \geq \bar{\psi}(x) - \underline{\psi}(x)\}.$$

Since $\bar{\psi}(x) = x = \underline{\psi}(x)$ for $x > \hat{y}$, $x_0 \leq \hat{y}$. Hence $\psi(y) = P_a\psi(y) = \hat{T}[y, L\psi(y)]$ for $\psi \in \{\bar{\psi}, \underline{\psi}\}$ and $y \leq x_0$. Moreover, by definition of x_0,

$$L\bar{\psi}(x_0) - L\underline{\psi}(x_0) < \bar{\psi}(x) - \underline{\psi}(x)$$

Clearly, $L\psi$ is continuous for $\psi \in \{\bar{\psi}, \underline{\psi}\}$. Hence, there exists $y \leq x_0$ such that $\bar{\psi}(y) - \underline{\psi}(y) \geq \bar{\psi}(x) - \underline{\psi}(x)$, and $L\bar{\psi}(y) - L\underline{\psi}(y) < \bar{\psi}(x) - \underline{\psi}(x)$. But since $\hat{T}$ is

1-Lipschitz in l,

$$\begin{aligned}\bar{\psi}(x) - \underline{\psi}(x) &> L\bar{\psi}(y) - L\underline{\psi}(y) \\ &\geq \hat{T}[y, L\bar{\psi}(y)] - \hat{T}[y, L\underline{\psi}(y)] \\ &= P_a\bar{\psi}(y) - P_a\underline{\psi}(y) \\ &= \bar{\psi}(y) - \underline{\psi}(y).\end{aligned}$$

This contradicts the definiton of y. The claim follows.

Let $\tilde{z} = \inf\{x \geq 0 : \hat{\sigma}_s(x) < 1\}$. Then, by construction of σ_a

$$w_{\sigma_a}(x) = \begin{cases} \bar{T}[1/2, x, Lw_{\sigma_a}(x)] & \text{for } x \in [\hat{z}, \hat{y}] \\ Lw_{\sigma_a}(x) - r\alpha x & \text{for } x \in [\tilde{z}, \hat{z}] \\ \bar{T}[1, x, Lw_{\sigma_a}(x)] & \text{for } x \leq \tilde{z}. \end{cases}$$

Moreover, for all $x \leq \tilde{z}$,

$$1 \geq \max \arg\max_a \bar{T}[a, x, Lw_{\sigma_a}(x)] \geq Q[x, Lw_{\sigma_a}(x)] = \sigma_a(x) = 1.$$

Hence, $w_{\sigma_a}(x) = \bar{T}[1, x, Lw_{\sigma_a}(x)] = \max_a \bar{T}[a, x, Lw_{\sigma_a}(x)]$. Then, σ_s is a fixed-point of P_a. Note that $\psi_\sigma \in C_a$ for all $\sigma \in \Sigma$. Then, clearly, $\psi \in C_a$. Then, from Claim 2, $\psi \leq P_a\psi$. For all $x \geq 0$ and $n \in \mathbb{N}$, let $\psi_n(x) = (P_a)^n\psi(x)$. Since P_a is order-preserving it follows that, viewing ψ_n as a function of $x \geq 0$ for all $n \in \mathbb{N}$, $(\psi_n)_n$ is increasing point-wise. Moreover, Since P_a maps C_a to C_a and $\psi \in C_a$, it follows that $\psi_n \in C_a$ for all $n \in \mathbb{N}$. In particular, $\psi_n \leq w_*$. Then, the sequence $(\psi_n)_n$ has a point-wise limit in C_a: call it $\hat{\psi}$. Then, clearly $\psi \leq \hat{\psi}$. Hence, it suffices to show that $\hat{\psi} = w_{\sigma_a}$. Since w_{σ_a} is the unique fixed-point of P_a in C_a, and $\hat{\psi} \in C_a$, it suffices to show that $\hat{\psi}$ is a fixed-point of P_a. To this end, note that $L\psi_n$ tends to $L\hat{\psi}$ point-wise. Then, result follows since $\hat{T}$ is continuous. □

Lemma 3. *Suppose that $F(s) = 1 - e^{-\lambda s}$. Let $\beta \in \mathbb{R}$, $\gamma \geq 0$, $\bar{x} > 0$, and let $w_0 \in C$ solve*

$$w(x) = \frac{r\beta x + \gamma Lw(x)}{r + \gamma} \tag{17}$$

at $x = \bar{x}$. *Let* $\hat{w} : [0, \infty) \to \mathbb{R}$ *be such that*

$$\hat{w}(x) = \begin{cases} \gamma \left[w_0(\bar{x}) - \beta \left(\bar{x} + \dfrac{\gamma}{\lambda r} \right) \right] e^{\frac{r\lambda}{r+\gamma}(x-\bar{x})} + \beta \left(x + \dfrac{\gamma}{\lambda r} \right) & \text{if } 0 \leq x \leq \bar{x} \\ w_0(x) & \text{if } x > \bar{x}. \end{cases}$$

Then, $\hat{w}$ *is the unique measurable* $w : [0, \infty] \to \mathbb{R}$ *such that* (17) *holds for* $0 \leq x \leq \bar{x}$, *and* $w(x) = w_0(x)$ *for* $x \geq \bar{x}$.

Proof of Lemma 3. Let $w : [0, \infty) \to \mathbb{R}$ be measurable and such that $w(x) = w_0(x)$ for $x > \bar{x}$. Then, for $0 \leq x \leq \bar{x}$, (17) is equivalent to

$$w(x) = \frac{r\beta x}{r+\gamma} + \frac{\gamma}{r+\gamma} \left(\int_x^{\bar{x}} w(y) \lambda e^{\lambda(x-y)} \mathrm{d}y + \int_{\bar{x}}^{\infty} w_0(y) \lambda e^{\lambda(x-y)} \mathrm{d}y \right). \tag{18}$$

Note that (18) is a Volterra equation of the second kind with exponential kernel. From Polyanin and Manzhirov (1998), (18) admits a unique solution $w_1 : [0, \bar{x}] \to \mathbb{R}$ and, moreover, w_1 is continuous. Let $\tilde{w} : [0, \infty) \to \mathbb{R}$ be given by $\tilde{w}(x) = w_1(x)$ for $x \in [0, \bar{x}]$, and $\tilde{w}(x) = w_0(x)$ for $x > \bar{x}$. Then $\tilde{w} \in C$ and, moreover, $\tilde{w}$ is the unique $w \in C$ such that (a) (17) holds for all $0 \leq x \leq \bar{x}$ and (b) $w(x) = w_0(x)$ for $x > \bar{x}$. Hence, it remains to show that $\tilde{w}(x) = \hat{w}(x)$ for all $0 \leq x \leq \bar{x}$.

Note that, since $F(s) = 1 - e^{-\lambda s}$, $L\tilde{w}(x) = \int_x^{\infty} \tilde{w}(y) \lambda e^{\lambda(x-y)} \mathrm{d}y$. Then, since $\tilde{w}$ is continuous, $L\tilde{w}$ is differentiable on $(0, \bar{x})$ with

$$L\tilde{w}'(x) = \lambda [L\tilde{w}(x) - \tilde{w}(x)].$$

Note that (17) is equivalent to

$$r[\tilde{w}(x) - \beta x] = a[L\tilde{w}(x) - \tilde{w}(x)].$$

From (17), since $L\tilde{w}$ is differentiable on $(0, \bar{x})$, then so is $\tilde{w}$ and

$$\begin{aligned}(r+\gamma)\tilde{w}'(x) &= r\beta + \gamma L\tilde{w}'(x) \\ &= r\beta + \lambda\gamma[L\tilde{w}(x) - \tilde{w}(x)] \\ \frac{r+\gamma}{r}\tilde{w}'(x) &= \beta + \lambda[\tilde{w}(x) - \beta x] \\ &= \lambda\tilde{w}(x) + \beta(1-\lambda x).\end{aligned}$$

Since this holds for all $x \in (0, \bar{x})$, it is clear that $\tilde{w}(x) = \hat{w}(x)$ for all $x \leq \bar{x}$. □

Proof of Claim 1. Since Agent 2 exerts no effort, Agent 1 is faced with the single-agent problem. From Theorem 1 of Chapter 2, some decreasing Markov strategy σ^1 is a best-response for Agent 1. Since effort is linear, it is clear that σ^1 is bang-bang. Note that $\hat{y}$ matches the definition of $\hat{x}$, in Section 3 of Chapter 2. Hence Agent 1 has no profitable deviation.

From Section 3 of Chapter 2, the Bellman equation for the value function of Agent 2 is

$$rv^2(x) = \max_{a\in[0,1]} r[b(x) - a\alpha(x)] + (1+a)[Lv^2(x) - v^2(x)]$$

Hence, Agent 2 has no profitable deviation if, for all $x \geq x_0$,

$$r\alpha x \geq Lv^2_{\hat{\sigma}}(x) - v^2_{\hat{\sigma}}(x). \tag{19}$$

Note that $v_{\hat{\sigma}}(x) = x$ for $x > \hat{y}$. Since $\hat{\sigma}^1(x) + \hat{\sigma}^2(x) = 1$ for all $x \leq \hat{y}$, $v^2_{\hat{\sigma}}$ solves (17) with $\gamma = \beta = 1$ at each $x \leq \hat{y}$. Moreover

$$\begin{aligned}v^2_{\hat{\sigma}}(\hat{y}) &= \frac{r\hat{y} + L\hat{y}}{r+1} \\ &= \hat{y} + \frac{1/\lambda}{r+1} \\ \Rightarrow v^2_{\hat{\sigma}}(\hat{y}) - \beta\left(\hat{y} + \frac{\gamma}{\lambda r}\right) &= -\frac{1}{\lambda r(r+1)}.\end{aligned}$$

Then, from Lemma 3, for $x \leq \hat{y}$,

$$v_{\hat{\sigma}}^2(x) = -\frac{1}{\lambda r(r+1)} e^{\frac{r\lambda x - 1/\alpha}{1+r}} + x + \frac{1}{\lambda r}.$$

Hence

$$\begin{aligned}\lambda[Lv_{\hat{\sigma}}^2(x) - v_{\hat{\sigma}}^2(x)] &= \lambda r[v_{\hat{\sigma}}^2(x) - x] \\ &= 1 - \frac{1}{r+1} e^{\frac{r\lambda x - 1/\alpha}{1+r}}.\end{aligned}$$

Hence $Lv_{\hat{\sigma}}^2(x) - v_{\hat{\sigma}}^2(x)$ is decreasing on $[0, \hat{y}]$. Therefore, (19) holds for all $x \in (\hat{z}, \hat{y}]$ if and only if it holds at $x = \hat{z}$. Note that, for $x \in (\hat{z}, \hat{y}]$,

$$\begin{aligned}\lambda[Lw_a(x) - w_a(x)] &= \lambda r[w_a(x) - (1 - \alpha/2)x] \\ &= 1 - \frac{\alpha}{2} + \frac{1}{2}\left[\alpha - \frac{1}{r+1}\right] e^{\frac{r\lambda x - 1/\alpha}{1+r}}.\end{aligned}$$

Let

$$\begin{aligned}\delta &= \exp\left\{\frac{r\lambda\hat{z} - 1/\alpha}{1+r}\right\} \\ \delta' &= \frac{\alpha(1+r)}{\alpha(1+r)+1}.\end{aligned}$$

Since $r\alpha\hat{z} = Lw_a(\hat{z}) - w_a(\hat{z})$,

$$\begin{aligned}\lambda\left\{r\alpha\hat{z} - \left[Lv_{\hat{\sigma}}^2(\hat{z}) - v_{\hat{\sigma}}^2(\hat{z})\right]\right\} &= 1 - \frac{\alpha}{2} + \frac{1}{2}\left[\alpha - \frac{1}{r+1}\right]\delta - 1 + \frac{\delta}{1+r} \\ &= \frac{1}{2}\left[\left(\alpha + \frac{1}{1+r}\right)\delta - \alpha\right].\end{aligned}$$

Hence, (19) holds at $x = \hat{z}$ if and only if $\delta' \leq \delta$. Let $g : (0, 1] \to \mathbb{R}$ be given by

$$g(d) = 2\alpha(1+r)\log d + \alpha - \left[\frac{1}{r+1} - \alpha\right] d.$$

Note that $g(\delta) = 0$ since $r\alpha\hat{z} = Lw_a(\hat{z}) - w_a(\hat{z})$. Moreover, g is strictly increasing.[15] Hence, $\delta' \leq \delta$ is equivalent to $g(\delta') \leq 0$. Finally, $g(\delta') \leq 0$ is equivalent to (2).

15. Indeed, $g'(d) = 2\alpha(1+r)/d + 1/(1+r) - \alpha \geq 0$ since $d \leq 1$.

Result follows. □

Proof of Claim 4. Fix $x \geq 0$ and note that $\hat{T}(x,l) = \max\{g_1(l), \min\{g_2(l), g_3(l)\}\}$ where $g_i : [0,\infty) \to \mathbb{R}$ is continuously differentiable and 1-Lipschitz for all $i = 1,2,3$. Moreover, there are $0 \leq l_1 \leq l_2 \leq \infty$ such that $\hat{T}(x,l) = g_1(l)$ for all $l < l_1$, $\hat{T}(x,l) = g_2(l)$ for all $l_1 < l < l_2$, and $\hat{T}(x,l) = g_3(l)$ for $l > l_2$. Then, for all $l \in (0,\infty) \setminus \{l_1, l_2\}$, $\hat{T}(x,l)$ is continuously differentiable at l with $\hat{T}_2(x,l) \in \{g_i'(l)\}_{i=1}^2$. Result follows. □

5.2 Proofs for Section 3

Proof of Lemma 2. To prove Part 1 is equivalent to Part 2, note that, if the opponents of Player i play σ_s, the payoff maximisation problem of Player i is a dynamic programming problem with state (k_t^i, x_t) and control (a_t^i, d_t^i).[16] Then, given $x_0 \geq 0$, σ_s is a PBE if and only if, for any $k_t^i > x_t \geq x_0$, it is optimal for Player i to disclose $d_t^i = k_t^i$, given that σ_s is played by all agents (including Player i) as soon as the state changes.

Note that, in this setting, disclosing any $d_t^i \in (x_t, k_t^i]$ induces the same payoffs, since Player i discloses fully as soon as the state changes. Hence, σ_s is a PBE if and only if, for all $k_t^i > x_t$, disclosing $d_t^i = k_t^i$ is (weakly) better for Player i than disclosing nothing until the state changes. Since time plays no role, σ_s is a PBE if and only if disclosing $d_t^i = k_t^i$ immediately is better for Player i than disclosing $d_t^i = k_t^i$ after a small time period $\mathrm{d}t$. This proves that Part 1 is equivalent to Part 2.

It remains to prove that Part 2 is equivalent to Part 3. Dropping the subscripts, the value to Player i from disclosing immediately is $w_s(k)$, whereas the value from disclosing after $\mathrm{d}t$ is

$$\delta(k,x) = \max_{a\in[0,1]} r[b(x) - c(a,x)]\mathrm{d}t + (1 - r\mathrm{d}t)\Bigg[a\mathrm{d}tLw_s(k) + (N-1)\sigma_s(x)\mathrm{d}t\int_k^\infty w_s(y)\mathrm{d}F(y-x) + (1 - \{a + (N-1)\sigma(x)[1 - F(k-x)]\}\mathrm{d}t)w_s(k)\Bigg].$$

16. In particular, given the state, time plays no role.

Hence, the net value of delayed disclosure has the same sign as

$$\Delta(k,x) = \frac{\delta(k,x) - w_s(k)}{\mathrm{d}t}.$$

Neglecting terms in $\mathrm{d}t^2$ implies that $\Delta(k,x)$ satisfies (5). Result follows. □

Proof of Proposition 3. As noted in Section 3 of Chapter 2, effort in σ_s ceases at the least value $y \geq 0$ such that

$$Lb(y) - b(y) = rc_1(0,y).$$

Substituting $b(x) = x$, and $c_1(0,y) = \alpha y$, shows that $\hat{y} = \mu_F/(r\alpha)$ is the unique solution. Let

$$\tilde{y} = \inf\{x \geq 0 : \sigma_s(x) < 1\}.$$

To show that (9) holds, note that, for $x \in [\tilde{y}, \hat{y}]$, the first-order condition for optimal effort implies that

$$\begin{aligned} w_s(x) &= Lw_s(x) - r\alpha x \\ \Rightarrow w_s'(x) &= Lw_s'(x) - r\alpha. \end{aligned} \tag{20}$$

Since $w_s'(x) = b'(x) = 1$ for $x > \hat{y}$, iterating the expression for $w_s'(x)$ implies that

$$w_s'(x) = 1 - r\alpha g_F(\hat{y} - x). \tag{21}$$

From Theorem 2 in Chapter 2, w_s is continuous. Hence $w_s(x) = x$ for $x \geq \hat{y}$ and, for $x \in [\tilde{y}, \hat{y}]$,

$$\begin{aligned} w_s(x) &= w_s\hat{y} - \int_x^{\hat{y}} w_s'(z)\mathrm{d}z \\ &= \hat{y} - \int_x^{\hat{y}} 1 - r\alpha g_F(\hat{y} - z)\mathrm{d}z. \end{aligned}$$

Equation (9) follows. To derive (8), note that

$$rw_s(x) = rx[1 - \alpha\sigma_s(x)] + N\sigma_s(x)[Lw_s(x) - w_s(x)].$$

Then, for $x \in [\tilde{y}, \hat{y}]$, (20) implies that

$$\begin{aligned} & rw_s(x) = rx[1 - \alpha\sigma_s(x)] + rN\sigma_s(x)\alpha x \\ \Leftrightarrow \quad & \sigma_s(x) = \frac{1}{N-1}\frac{w_s(x) - x}{\alpha x}. \end{aligned}$$

Equation (8) follows by substituting (9) for $w_s(x)$. Finally, $\tilde{y}$ solves (7) since σ_s is continuous at $\tilde{y}$, and is given by (8) on $[\tilde{y}, \hat{y}]$. □

Lemma 4. *Suppose that F is continuous. Then, for all $x \geq 0$*

$$\lim_{\gamma \to 0} g_F\left(\frac{\mu_F}{\gamma} - x\right) - \frac{1}{\gamma} = \frac{\mu_{F^2}}{2\mu_F^2} - \frac{x}{\mu_F}.$$

Proof of Lemma 4. Let $S_1, S_2, \ldots$ be i.i.d. draws from F. For $m \in \mathbb{N}$, let $X(m) = \sum_{i=1}^{m} S_i$. For all $z \geq 0$, let M_z be the smallest $m \in \mathbb{N}$ such that $X(m) > z$. Then, since F is continuous,

$$\begin{aligned} & g_F(z) \\ & \quad = 1 + \max\{m \in \mathbb{N} : X(m) \leq z\} \\ & \quad = \mathbb{E}[M_z]. \end{aligned}$$

From Wald's equation

$$\mu_F \mathbb{E}[M_z] = \mathbb{E}[X(M_z)].$$

Moreover, since F is continuous, the renewal theorem implies that

$$\lim_{z \to \infty} \mathbb{E}[X(M_z)] - z = \frac{\mu_{F^2}}{2\mu_F}.$$

Hence

$$\begin{aligned} g_F\left(\frac{\mu_F}{\gamma} - x\right) - \frac{1}{\gamma} &= \frac{1}{\mu_F}\left[\mu_F g_F\left(\frac{\mu_F}{\gamma} - x\right) - \frac{\mu_F}{\gamma} + x\right] - \frac{x}{\mu_F} \\ &= \frac{1}{\mu_F}\left[\mu_F \mathbb{E}\left[M_{\mu_F/\gamma - \tilde{x}}\right] - \frac{\mu_F}{\gamma} + x\right] - \frac{x}{\mu_F} \\ &= \frac{1}{\mu_F}\left[\mathbb{E}\left[X(M_{\mu_F/\gamma - x})\right] - \frac{\mu_F}{\gamma} + x\right] - \frac{x}{\mu_F}. \end{aligned}$$

Result follows by letting γ tend to 0. □

Proof of Corollary 1. The first part follows from (7). From (8), w_s is independent of N on $[\tilde{y}, \infty)$. Moreover, $w_s'(x) = 1$ for $x > \hat{y}$ and (18) holds on $(\tilde{y}, \hat{y})$. Since g_F is positive and increasing, it follows that w_s is convex on $[\tilde{y}, \infty)$. Moreover, w_s has a unique minimum unless there exist $z_1 \neq z_2$ such that $g_F(z_i) = r\alpha$ for $i = 1, 2$. Clearly, this holds for at most countably many values of $r\alpha$. Hence, in general, w_s has a unique minimum (call it $\tilde{x}$) on $[\tilde{y}, \infty)$. Note that w_s' is bounded on $[0, \tilde{y}]$ as N diverges. Hence, for N large enough, w_s is minimised at $\tilde{x}$. The limit for $\tilde{x}$ follows from Lemma 4. □

Proof of Proposition 4. The first part follows from the reasoning in the paragraph before the statement of the proposition. To prove the second part, fix $x_0 > \tilde{x}_F$. Then, it suffices to show that (if α is large and, given α, N is large) $\Delta(x, k) \leq 0$ for all $x_0 \leq x \leq k$. Fix $x \geq x_0$, and note that $\Delta(x, x) = 0$. Then, it suffices to show that $\Delta(k, x)$ is decreasing in k on $[x, \infty)$. There are two cases.

Case 1: $k \leq \hat{y}$. Fix $\alpha > 0$, and suppose that N is large enough so that $x_0 \geq \tilde{y}$. Then, since $x \geq \tilde{y}$, $Lw(k) - w(k) = r\alpha k$ is increasing in k on $[x_0, \hat{y}]$. Moreover, as $\sigma_s(k) > 0$ for $k \leq \hat{y}$, $-rc(a, x) + a[Lw_s(k) - w_s(k)]$ is maximised by $a = 1$. It follows that $a = 1$ is also a maximand in the expression for $\Delta(k, x_0)$. Then, by the enveloppe theorem, the partial derivative of Δ with respect to k exists over $(x_0, \hat{y})$ and takes value

$$\begin{aligned}
&\Delta_1(k, x) = (Lw)'(k) - w'(k) - \{(N-1)\sigma(x_0)[1 - F(k - x_0)] + r\}w'(k) \\
&\frac{\Delta_1(k, x)}{r\alpha} = 1 - \left[\frac{1}{x_0}\int_{x_0}^{\hat{y}} g_F(\hat{y} - x)\mathrm{d}x[1 - F(k - x_0)] + r\right]\left[\frac{1}{r\alpha} - g_F(\hat{y} - k)\right]. \quad (22)
\end{aligned}$$

Then, it suffices to show that $\Delta_1(k, x) \leq 0$. Note that

$$\frac{\Delta_1(k, x)}{r\alpha} \leq 1 - r\left[\frac{1}{r\alpha} - g_F(\hat{y} - k)\right].$$

Let $\hat{k}$ be such that

$$1 - \frac{r}{\mu_F}\left(\hat{k} - \tilde{x}_F\right) < 0.$$

Then, from Lemma 4, $\Delta_1(\hat{k}, x) < 0$ for $\alpha > 0$ sufficiently small. Note that g_F is

increasing. Hence, for $\alpha > 0$ sufficiently small, $\Delta_1(k,x) < 0$ for all $k \geq \hat{k}$. Then, there exists $\hat{k}$ such that, for $k > \hat{k}$, $\Delta_1(k,x) < 0$ if α sufficiently small. Hence, suppose that $x \leq k \leq \hat{k}$. By assumption,

$$\frac{x_0}{\tilde{x}_F} - 1 > 0.$$

Then, since $k \geq x \geq x_0$, and $\hat{k} \geq k \geq x \geq x_0$, for α sufficiently small,

$$\Delta_1(k,x) \leq 1 - \left[\frac{1}{x_0}\int_{x_0}^{\hat{y}} g_F(\hat{y}-x)\mathrm{d}x\left[1 - F\left(\hat{k} - x_0\right)\right] + r\right]\left(\frac{x_0}{\tilde{x}_F} - 1\right).$$

Note that $\hat{y} = \mu_F/(r\alpha) \to \infty$ as $\alpha \to 0$. Then, for α sufficiently small, the integral is sufficiently large. Since F has unbounded support, $\left[1 - F\left(\hat{k} - x_0\right)\right] > 0$. Result follows.

Case 2: $k \geq \hat{y}$. Since $x_0 > \tilde{x}_F$, w_s is increasing on $[x, \infty)$. Then, the third and fourth terms in (5) are decreasing in k. Moreover, $Lw(k) - w(k) = r\hat{y}$ for $k \geq \hat{y}$. Hence, for $k \geq \hat{y}$, $\Delta(k,x) < \Delta(\hat{y},x)$. Therefore result follows from the first case. □

Chapter 4

The Lattice Structure of Preference Comparison

Abstract

Most comparisons of preferences have the structure of *single-crossing dominance*. We examine the lattice structure of single-crossing dominance, proving characterisation, existence and uniqueness results for minimum upper bounds of arbitrary sets of preferences. We apply these theorems to monotone comparative statics, ambiguity- and risk-aversion, social choice, and politically correct discourse.

1 Introduction

Comparisons of preferences are ubiquitous in economics: examples include 'more risk-averse/ambiguity-averse than' (in decision theory),[1] 'takes larger actions than' (in monotone comparative statics),[2] 'more delay-averse/impatient than' (in dynamic problems),[3] and 'more self-controlled than' (in models of temptation).[4] All of these preference comparisons, and most others besides, are special cases of *single-crossing dominance*, a general unified way of comparing preferences.

In this paper, we investigate the lattice structure of single-crossing dominance. In particular, in §2, we characterise the minimum upper bounds of arbitrary sets of preferences, and give necessary and sufficient conditions for their existence and uniqueness.

We use our theorems to derive new results in a variety of economic settings that feature some of the aforementioned preference comparisons. In §3, we develop a theory of

1. Yaari (1969), Epstein (1999) and Ghirardato and Marinacci (2002).
2. Topkis (1978), Milgrom and Shannon (1994) and LiCalzi and Veinott (1992).
3. Horowitz (1992) and Benoît and Ok (2007).
4. Gül and Pesendorfer (2001) and Dekel and Lipman (2008).

monotone comparative statics for the consensus of a group of agents as well as a theory of misspecification-robust comparative statics. In §4, we characterise a general class of maxmin preferences over acts as precisely minimum upper bounds with respect to 'more ambiguity-averse than'. In §5, we characterise when acceptable aggregation of individual preferences is possible in the presence of normative constraints. Additionally, alongside the general theory in §2, we study the political correctness, coherence and concordance of expressed views in a politically correct society.

1.1 Theory

The abstract environment consists of a non-empty set $\mathcal{X}$ of alternatives equipped with a partial order $\gtrsim$.[5] Write $\mathcal{P}$ for the set of all preferences (complete and transitive binary relations) on $\mathcal{X}$. One preference single-crossing dominates another iff the former is more disposed to favour $\gtrsim$-larger alternatives:

Definition 1. For two preferences $\succeq, \succeq' \in \mathcal{P}$, we write $\succeq' S \succeq$ iff for any pair $x \gtrsim y$ in $\mathcal{X}$, $x \succeq (\succ)\ y$ implies $x \succeq' (\succ')\ y$. We say that $\succeq'$ *single-crossing dominates* $\succeq$.[6]

Given a set of preferences $P \subseteq \mathcal{P}$, a preference $\succeq' \in \mathcal{P}$ is an *upper bound* of P iff $\succeq' S \succeq$ for every $\succeq \in P$, and a *minimum* upper bound iff in addition $\succeq'' S \succeq'$ for every (other) upper bound $\succeq''$ of P.[7] Intuitively, a minimum upper bound of P is a preference that likes large alternatives more than does any preference in P, but only just. Maximum lower bounds are defined analogously.

In §2, we study the *(complete) lattice structure* of the poset $(\mathcal{P}, S)$ by developing characterisation, existence and uniqueness results for minimum upper bounds. Our characterisation theorem (§2.1) characterises the minimum upper bounds of arbitrary sets $P \subseteq \mathcal{P}$ of preferences. Our existence theorem (§2.2) identifies the condition on $\gtrsim$, called *crown- and diamond-freeness,* that is necessary and sufficient for every $P \subseteq \mathcal{P}$ to possess

5. Definitions of standard concepts such as these are collected in appendix A.

6. This definition is from Milgrom and Shannon (1994), in the context of monotone comparative statics. This type of preference comparison has been used at least since Yaari's (1969) definition of 'more risk-averse than' (see §4.4). The term 'single-crossing' originates in early information economics, where the '(Spence–Mirrlees) single-crossing condition' plays an important role (Mirrlees 1971; Spence 1973).

7. One may similarly define a *minimal* upper bound as an upper bound $\succeq'$ such that there is no (other) upper bound $\succeq''$ with $\succeq' \not S \succeq'' S \succeq'$. We do not study minimal upper bounds in this paper.

a minimum upper bound. (In appendix I, we provide a graphical characterisation of crown- and diamond-freeness in terms of the *transitive reduction*.) Finally, our uniqueness proposition (§2.3) asserts that every set $P \subseteq \mathcal{P}$ has a *unique* minimum upper bound precisely if $\gtrsim$ is *complete*. We extend our results to maximum lower bounds in appendix F.

1.2 Applications

We employ our theorems to answer a broad range of economic questions.

Application to politically correct discourse (§2)**.** This application is developed in §2 alongside the general theory. There is a set $\mathcal{X}$ of social issues. Political correctness demands that some issues x be prioritied over others y; we write $x \gtrsim y$ when this is the case.

A set of agents have coherent (i.e. transitive) priorities, formally anti-symmetric preferences $P \subseteq \mathcal{P}$. An agent's view (priority) $\succeq \in P$ on a pair of issues $x, y \in \mathcal{X}$ with $x \gtrsim y$ is *politically correct* if $x \succ y$, and *politically incorrect* otherwise. A priority $\succeq'$ is *more politically correct than* another priority $\succeq$ iff the former holds a politically correct view whenever the latter does—that is, iff $\succeq' S \succeq$.

The agents meet and discuss, and leave the conversation with expressed views (priorites), formed (roughly) as follows. If an agent's view on a pair of issues is politically incorrect, but her conversational partner holds the politically correct view, then the agent self-censors by expressing the politically correct priority. Otherwise, she expresses her true view.

How politically correct will agents' expressed views be? When can we be assured that expressed views are coherent? And when is consensus assured? We answer these questions alongside the general theory in §2.

Application to monotone comparative statics (§3)**.** In monotone comparative statics, an agent chooses an alternative from a set $\mathcal{X} \subseteq \mathbf{R}$. Write $X(\succeq)$ for the set of optimal alternatives when the agent's preference is $\succeq \in \mathcal{P}$. The canonical result states that if $\succeq$ increases in the sense of S, then $X(\succeq)$ increases in the strong set order. We extend this result in two directions.

First, suppose that there is a group of agents with preferences $P \subseteq \mathcal{P}$. The *consensus* is the set of alternatives that they all find optimal: $C(P) = \bigcap_{\succeq \in P} X(\succeq)$. Using the existence theorem, we show that when the set P increases in the strong set order, so does the consensus $C(P)$.

Second, suppose that an analyst desires a comparative statics conclusion, but knows only that the agent's preference lies in a set $P \subseteq \mathcal{P}$. We show using the existence theorem that the set $X(P) \coloneqq \bigcup_{\succeq \in P} X(\succeq)$ of *possibly-optimal* alternatives increases in a natural weakening of the strong set order whenever P does.

Application to ambiguity- and risk-aversion (§4)**.** Consider the standard Savage framework with monetary prizes, and let $\mathcal{X}$ be a set of acts. One preference is *more ambiguity-averse than* another iff whenever the latter (strictly) prefers a constant act to a non-constant one, so does the latter.

Say that a set $P \subseteq \mathcal{P}$ of monotone and solvable preferences is a *maxmin representation* of a preference $\succeq^\star \in \mathcal{P}$ iff the latter is ordinally represented by $X \mapsto \inf_{\succeq \in P} c(\succeq, X)$, where $c(\succeq, X)$ denotes the certainty equivalent of an act X according to preference $\succeq$. Maxmin expected utility (Gilboa and Schmeidler 1989) is a (very) special case.

We characterise the maxmin class by proving that P is a maxmin representation of $\succeq^\star$ if, and only if, $\succeq^\star$ is a minimum upper bound of P with respect to 'more ambiguity-averse than'. A comparative statics result follows: when P increases in appropriate sense, $\succeq^\star$ becomes more ambiguity-averse. Another corollary is that every set of monotone and solvable preferences over acts possesses exactly one minimum upper bound.

These results carry over to *risk*-aversion.

Application to social choice (§5)**.** A group of individuals, each with preferences over a set $\mathcal{X}$ of social alternatives, must aggregate their preferences into a social preference.

For certain pairs $x, y \in \mathcal{X}$ of alternatives, it is normatively acceptable to rank y above x only with the consent of every individual—we write $x \gtrsim y$ when this is the case. A SWF is *acceptable* iff it respects this constraint and satisfies a suitable Pareto criterion.

We prove an (im)possibility theorem that gives the necessary and sufficient condition on $\gtrsim$ for acceptable SWF to exist. The conditions are stringent—they imply crown- and

diamond-freeness—yielding a negative conclusion akin to Sen's (1970) 'impossibility of a Paretian liberal'. But there are good news, too: we exhibit some natural normative constraints for which acceptable aggregation is possible according to our theorem.

1.3 Related literature

Our work relates to the combinatorics literature on permutation lattices (Bennett and Birkhoff 1994; Markowsky 1991; Duquenne and Cherfouh 1994). Here the alternatives are $\mathcal{X} = \{1, \ldots, n\}$ for some $n \in \mathbf{N}$, and the partial order $\gtrsim$ is the ordinary inequality. This means in particular that $\gtrsim$ is complete and that $\mathcal{X}$ is finite.

In this context, anti-symmetric (i.e. never-indifferent) preferences may be thought of as *permutations*, and single-crossing dominance is known as the *weak order*.[8] It has been known since Guilbaud and Rosenstiehl (1963) and Yanagimoto and Okamoto (1969) that the set of all permutations equipped with the weak order is a (complete) lattice. Châtel, Pilaud and Pons (2018) extend this result to other sets of binary relations on $\mathcal{X}$. Our uniqueness proposition is a result along these lines.

Since this literature assumes that $\gtrsim$ is complete, it certainly contains no analogue of our existence theorem. We are not aware of any analogue of our characterisation theorem, either. Besides avoiding the restrictive assumption that $\gtrsim$ is complete, we differ from this literature by allowing for preferences with indifferences and by permitting $\mathcal{X}$ to be of unrestricted cardinality.

Crown- and diamond-freeness are standard concepts in combinatorics; the former is studied by e.g. Lu (2014), and the latter by Griggs, Li and Lu (2012). Truss (1996), Ball, Pultr and Sichler (2006) and Gray and Truss (2009) study crown- and diamond-freeness together.

We discuss literature relevant to the applications in their respective sections.

8. Also called the 'right permutahedron order', or the 'weak Bruhat order'.

2 Theory

In this section, we develop our general results about the lattice structure of single-crossing: the characterisation theorem (§2.1), the existence theorem (§2.2) and the uniqueness proposition (§2.3). Recall the abstract environment and basic definitions from §1.1.

To motivate and illustrate the results, we introduce the following model of public discourse in the presence of political correctness.

Application to politically correct discourse. In a society with norms of political correctness, people frequently self-censor rather than express their politically incorrect views publicly. We ask how much self-censorship occurs, when self-censorship may lead to incoherent expressed views, and when self-censorship is so severe that it leads to a consensus of expressed opinions.

Loury (1994) and Morris (2001) also study discourse under political correctness. Their framework and questions differ greatly from ours: Morris (2001) considers a game of cheap-talk communication in which the sender knows that the receiver entertains doubts about whether she is biased, and studies how this impedes information transmission. Loury (1994) tells a similar story without a formal model.

There is a set $\mathcal{X}$ of social issues. Some issues x have higher political priority than others y, meaning that it is politically incorrect to express the view that y should be prioritised over x; we write $x \gtrsim y$ when this is the case. Assume that $\gtrsim$ is a partial order.

For instance, the political issues could include economic justice for ethnic minorities (x), for women (y), and for ethnic-majority men (z). In this case, political correctness may demand that ethnic minorities and women be prioritied over ethnic-majority men ($x \gtrsim z$ and $y \gtrsim z$), while remaining silent on the relative importance of economic justice for ethnic minorities and for women (x, y are $\gtrsim$-incomparable).

A *priority* is a complete and (for simplicity) anti-symmetric (i.e. never-indifferent) binary relation $\succeq$ on $\mathcal{X}$. '$x \succ y$' expresses the view that society ought to prioritise issue x over issue y. A priority is *coherent* iff it is transitive.

Say that a priority $\succeq'$ is *more politically correct* than another priority $\succeq$ iff for any pair $x, y \in \mathcal{X}$ with $x \gtrsim y$, $x \succ y$ implies $x \succ' y$. Observe that this is precisely single-crossing

dominance, extended to possibly intransitive relations.[9]

There are two agents, whose true views are described by coherent priorities $P = \{\succeq_1, \succeq_2\}$. (The analysis extends straightforwardly to the case of many agents.) The agents converse about about what social priorities over $\mathcal{X}$ ought to be, and leave the conversation with *expressed* priorities $\succeq'_1$ and $\succeq'_2$.

An agent is willing to express to her politically incorrect view on a pair of issues $x, y \in \mathcal{X}$ only if her interlocutor holds the same view; otherwise, she self-censors. In particular, consider issues $x \gtrsim y$. If either agent holds the politically correct view that x should be prioritised over y, then both express the politically correct view. (Schematically, if either $x \succ_1 y$ or $x \succ_2 y$, then $x \succ'_1 y$ and $x \succ'_2 y$.) There is a caveat: if $x_1 \gtrsim \cdots \gtrsim x_K$ and some agent holds the politically correct view on each pair x_k, x_{k+1}, then agents self-censor on x_1, x_K by expressing the view that x_1 should be prioritised over x_K, even if both hold the politically incorrect view on x_1, x_K. There is at least one priority that satisfies these constraints. If there are several, then we assume that agents express a coherent priority if possible.

To illustrate, suppose that there are three issues $\mathcal{X} = \{x, y, z\}$, and that political correctness demands that $x > y > z$. If agents's views are $z \succ_1 x \succ_1 y$ and $y \succ_2 z \succ_2 x$, then both express the view $x \succ'_i y$ since $x \gtrsim y$ and $x \succ_1 y$, and both express $y \succ'_i z$ since $y \gtrsim z$ and $y \succ_2 z$. They therefore also express the politically correct view $x \succ'_i z$, even though both agents' true views on x, z are politically incorrect (viz. $z \succ_i x$).

How much self-censorship will occur? (§2.1: by the characterisation theorem, both agents' expressed priorities will be more politically correct than either agent's true view, but no more politically correct than that.) Do agents express coherent (i.e. transitive) priorities, no matter what their (coherent) true views $P = \{\succeq_1, \succeq_2\}$? (§2.2: yes iff $\gtrsim$ is *crown- and diamond-free,* by the existence theorem.) Is there expressed consensus, no matter what the true views P? (§2.3: yes iff $\gtrsim$ is complete, by the uniqueness proposition.)

9. Recall that we are considering only anti-symmetric relations.

2.1 Characterisation of minimum upper bounds

In this section, we characterise the minimum upper bounds of an arbitrary set $P \subseteq \mathcal{P}$ of preferences. Our characterisation is in terms of *P-chains*, defined as follows.

Definition 2. For a set $P \subseteq \mathcal{P}$ of preferences and two alternatives $x \gtrsim y$ in $\mathcal{X}$, a *P-chain* from x to y is a finite sequence $(w_k)_{k=1}^{K}$ in $\mathcal{X}$ such that

(i) $w_1 = x$ and $w_K = y$,

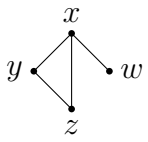
(ii) for every $k < K$, $w_k \gtrsim w_{k+1}$, and

(iii) for every $k < K$, $w_k \succeq w_{k+1}$ for some $\succeq \in P$.

A *strict P-chain* is a P-chain with $w_k \succ w_{k+1}$ for some $k < K$ and $\succeq \in P$.

In a word, P-chain is a $\gtrsim$-decreasing sequence of alternatives along which, at each juncture (each $k < K$), some preference in P prefers the previous ($\gtrsim$-larger) alternative to the subsequent ($\gtrsim$-smaller) one. Note that a P-chain of length $K \geq 3$ is simply the concatenation of $K - 1$ P-chains of length 2.

Example 1. Consider $\mathcal{X} = \{x, y, z, w\}$, with $\gtrsim$ such that $x > w$ and $x > y > z$ (so $x > z$), and w, y and w, z are incomparable. The partial order $\gtrsim$ may be depicted graphically as



In this (standard) diagram, there is an edge ('link') from *a down* to b iff $a \gtrsim b$. The absence of an edge therefore represents $\gtrsim$-incomparability. We will use diagrams of this sort throughout.

Consider $P = \{\succeq_1, \succeq_2\}$, where

$$z \succ_1 w \succ_1 x \succ_1 y \quad \text{and} \quad y \succ_2 z \succ_2 w \succ_2 x.$$

Since $x \gtrsim y$ and $x \succ_1 y$, (x, y) is a strict P-chain. (y, z) is similarly a strict P-chain. It

follows that there is a strict P-chain from x to z, namely (x, y, z). Note, however, that (x, z) is *not* a P-chain, since neither preference ranks x above z.

Although $x \gtrsim w$, there is no P-chain from x to w: the only candidate is (x, w), and it fails to be a P-chain because neither preference favours x over w.

We begin by characterising *upper bounds* (minimum or not) in terms of P-chains. An upper bound of $P \subseteq \mathcal{P}$ is precisely a preference that (strictly) prefers a larger alternative to a smaller one whenever there is a (strict) P-chain between them:

UB characterisation lemma. For a preference $\succeq' \in \mathcal{P}$ and a set $P \subseteq \mathcal{P}$ of preferences, the following are equivalent:

(1) $\succeq'$ is an upper bound of P.

(2) $\succeq'$ satisfies: for any $\gtrsim$-comparable $x, y \in \mathcal{X}$, wlog $x \gtrsim y$,

(i) $x \succeq' y$ if there is a P-chain from x to y, and

(ii) $y \succeq' x$ only if there is no strict P-chain from x to y.

Example 1 (continued)**.** The P-chains, all of them strict, are (x, y), (y, z) and (x, y, z). Thus by the UB characterisation lemma, a preference $\succeq' \in \mathcal{P}$ is an upper bound of P iff $x \succ' y \succ' z$ (and $x \succ' z$). Thus,

$$x \succ'_a y \succ'_a z \succ'_a w \quad \text{and} \quad w \succ'_b x \succ'_b y \succ'_b z$$

are both upper bounds.

Proof. (2) implies (1): Let $\succeq' \in \mathcal{P}$ satisfy condition (2); we wish to show that $\succeq' S \succeq$ for any $\succeq \in P$. To that end, fix a $\gtrsim$-comparable pair $x, y \in \mathcal{X}$, wlog $x \gtrsim y$, and suppose that $x \succeq (\succ) y$ for some $\succeq \in P$; we must show that $x \succeq' (\succ') y$. This is immediate since (x, y) is a (strict) P-chain.

(1) implies (2): Let $\succeq'$ be an upper bound of P. Fix a $\gtrsim$-comparable pair $x, y \in \mathcal{X}$, wlog $x \gtrsim y$. We must show that if there is a (strict) P-chain from x to y, then $x \succeq' (\succ') y$.

Suppose that there exists a P-chain $(w_k)_{k=1}^K$ from x to y. For each $k < K$, we have $w_k \gtrsim w_{k+1}$ as well as $w_k \succeq w_{k+1}$ for some $\succeq \in P$. Since $\succeq'$ is an upper bound of P, it must be that $w_k \succeq' w_{k+1}$ for each $k < K$. Since $\succeq'$ is transitive (because it lives in $\mathcal{P}$), it follows that $x \succeq' y$.

Suppose that there exists a strict P-chain $(w_k)_{k=1}^K$ from x to y. As in the weak case, it must be that $w_k \succeq' w_{k+1}$ for each $k < K$. Moreover, since the P-chain is strict, we have $w_k \succ w_{k+1}$ for some $k < K$ and $\succeq \in P$; hence $w_k \succ' w_{k+1}$ since $\succeq'$ is an upper bound of P. It follows by transitivity of $\succeq'$ that $x \succ' y$. ■

The UB characterisation lemma says that an upper bound must have a (strict) 'upward' preference whenever there is a (strict) P-chain. Our characterisation theorem says that the *minimum* upper bounds are those which have a (strict) 'upward' preference *only* when there is a (strict) P-chain:

Characterisation theorem. For a preference $\succeq^\star \in \mathcal{P}$ and a set $P \subseteq \mathcal{P}$ of preferences, the following are equivalent:

(1) $\succeq^\star$ is a minimum upper bound of P.

(2) $\succeq^\star$ satisfies: for any $\gtrsim$-comparable $x, y \in \mathcal{X}$, wlog $x \gtrsim y$,

($\star$) $x \succeq^\star y$ iff there is a P-chain from x to y, and

($\star\star$) $y \succeq^\star x$ iff there is no strict P-chain from x to y.

The analogous characterisation for maximum lower bounds is given in appendix F.

Example 1 (continued)**.** A minimum upper bound $\succeq^\star \in \mathcal{P}$ must satisfy $x \succ^\star y \succ^\star z$ (and $x \succ^\star z$) since it is an upper bound. Since $x \gtrsim w$ but there is no P-chain from x to w, minimumhood requires that $w \succ^\star x$ by the characterisation theorem. In sum, $\succeq^\star \in \mathcal{P}$ is a minimum upper bound iff

$$w \succ^\star x \succ^\star y \succ^\star z.$$

(There is clearly exactly one such preference.)

One direction of the proof is straightforward:

Proof that (2) implies (1). Fix a subset P of $\mathcal{P}$ and a $\succeq^\star \in \mathcal{P}$ that satisfies $(\star)$–$(\star\star)$. It is immediate from the UB characterisation lemma that $\succeq^\star$ is an upper bound of P.

To see that $\succeq^\star$ is a minimum of the upper bounds of P, let $\succeq'$ be any upper bound of P. Fix a $\gtrsim$-comparable pair $x, y \in \mathcal{X}$, wlog $x \gtrsim y$, and suppose that $x \succeq^\star (\succ^\star) y$. By property $(\star)$ (property $(\star\star)$), there must be a (strict) P-chain from x to y. Since $\succeq'$ is an upper bound of P, it follows by the UB characterisation lemma that $x \succeq' (\succ') y$. Since $x, y \in \mathcal{X}$ were arbitrary, this establishes that $\succeq' S \succeq^\star$. ■

The other direction requires the following lemma, whose (rather involved) proof is given in appendix C (p. 159).

Lemma 1. Let P be a set of preferences, and let $x, y \in \mathcal{X}$ be $\gtrsim$-comparable, wlog $x \gtrsim y$. If there is no (strict) P-chain from x to y, then there exists an upper bound $\succeq''$ of P with $x \not\succeq'' (\not\succ'') y$.

Proof that (1) implies (2). Fix a set $P \subseteq \mathcal{P}$ of preferences and a preference $\succeq' \in \mathcal{P}$. We will prove the contra-positive: if $\succeq'$ violates $(\star)$–$(\star\star)$, then it cannot be a minimum upper bound of P. If a preference $\succeq'$ violates the 'if' part of $(\star)$ or the 'only if' part of $(\star\star)$, then it fails to be an upper bound of P by the UB characterisation lemma.

Suppose that a preference $\succeq'$ violates the 'only if' part of $(\star)$ (the 'if' part of $(\star\star)$). In particular, there are $\gtrsim$-comparable $x, y \in \mathcal{X}$, wlog $x \gtrsim y$, such that there is no (strict) P-chain from x to y, and yet $x \succeq' (\succ') y$. By Lemma 1, there is an upper bound $\succeq''$ of P such that $x \not\succeq'' (\not\succ'') y$. Then $\succeq'' \not\supseteq \succeq'$, so $\succeq'$ fails to be a minimum of the upper bounds of P. ■

Application to politically correct discourse (continued)**.** The characterisation theorem permits us to show that agents will express priorities that are more politically correct than either agent's true view, but no more politically correct than that. That is, their expressed views $\succeq_1'$ and $\succeq_2'$ are both minimum upper bounds (with respect to 'more politically correct than') of $P = \{\succeq_1, \succeq_2\}$.

Suppose that $x > y$ and that there is a (necessarily strict) P-chain $(w_k)_{k=1}^K$ from x to

y.[10] For each $k < K$, since $w_k \gtrsim w_{k+1}$ and one of the agents holds the politically correct view $w_k \succ_i w_{k+1}$, both agents' expressed priorities favour w_k over w_{k+1}. Both expressed views $\succeq_1'$ and $\succeq_2'$ therefore prioritise x over y.

Suppose on the other hand that $x > y$ that there is no P-chain from x to y. Then both agents express their politically incorrect view that y should be prioritsed over x.

Since $x, y \in \mathcal{X}$ were arbitrary, this shows that $\succeq_1'$ and $\succeq_2'$ satisfy properties $(\star)$–$(\star\star)$. Thus, by the characterisation theorem, the expressed views $\succeq_1'$ and $\succeq_2'$ are minimum upper bounds of the agents' true views $P = \{\succeq_1, \succeq_2\}$, in the extended sense that does not require them to be coherent (i.e. to live in $\mathcal{P}$).

2.2 Existence of minimum upper bounds

In this section, we provide a necessary and sufficient condition on $\gtrsim$ for minimum upper bounds to exist for every set of preferences. This condition rules out two kinds of special subposet: *crowns* and *diamonds.*

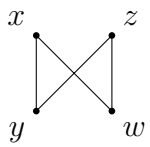
In §2.2.1, we define crowns and show that $\gtrsim$ must be free of them if every set of preferences is to possess a minimum upper bound. In §2.2.2, we do the same for diamonds. In §2.2.3, we give the existence theorem, which asserts in addition that crown- and diamond-freeness of $\gtrsim$ is *sufficient* for every set of preferences to possess a minimum upper bound. In appendix I, we provide an alternative characterisation of crown- and diamond-freeness in terms of the transitive reduction.

2.2.1 Crowns

The following example identifies a $\gtrsim$ for which a certain $P \subseteq \mathcal{P}$ admits no minimum upper bound:

10. All P-chains are strict because we assumed (for simplicity) that $\succeq_1$ and $\succeq_2$ are anti-symmetric (never indifferent).

Example 2. Consider $\mathcal{X} = \{x, y, z, w\}$ with the following partial order $\gtrsim$:



That is, each of x and z $\gtrsim$-dominate each of y and w, but x, z and are $\gtrsim$-incomparable, as are y, w. Let $P = \{\succeq_1, \succeq_2\} \subseteq \mathcal{P}$, where

$$w \succ_1 x \succ_1 y \succ_1 z \quad \text{and} \quad y \succ_2 z \succ_2 w \succ_2 x.$$

$x \gtrsim y$ and $z \gtrsim w$, and there is a strict P-chain from x to y and from z to w. On the other hand, $x \gtrsim w$ and $z \gtrsim y$, but there is no P-chain from x to w or from z to y. So by the characterisation theorem (p. 127), a minimum upper bound $\succeq^\star$ of P must have

$$x \succ^\star y \succ^\star z \succ^\star w \succ^\star x.$$

Such a $\succeq^\star$ cannot be transitive, so cannot live in $\mathcal{P}$. It follows that no minimum upper bound exists.

The feature of $\gtrsim$ in Example 2 that causes existence to fail is that is contains a *crown*, defined as follows.

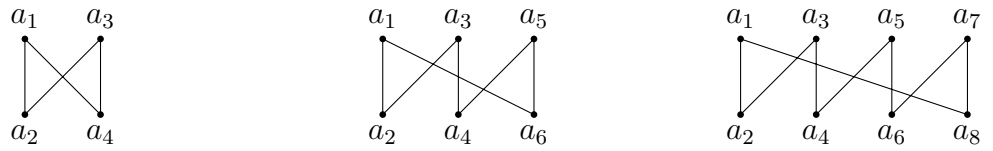
Definition 3. Let $\geq$ be a binary relation on a set $\mathcal{A}$. For $K \geq 4$ even, a *K-crown* is a sequence $(a_k)_{k=1}^K$ in $\mathcal{A}$ such that non-adjacent $a_k, a_{k'}$ are $\geq$-incomparable,[11] and $a_{k-1} > a_k < a_{k+1}$ for each $1 < k \leq K$ even (where $a_{K+1} \coloneqq a_1$ by convention). A *crown* is a K-crown for some $K \geq 4$ even. $\geq$ is *crown-free* iff it contains no crowns.

Some crowns are drawn in Figure 1.

Example 2 (continued)**.** $\gtrsim$ contains the 4-crown (x, y, z, w), so fails to be crown-free.

Crown-freeness rules out a specific form of incompleteness. A strong sufficient condition is completeness. A weaker sufficient condition is that the comparability relation $\gtrless$ be

11. Schematically: for $1 \leq k < k+2 \leq k' \leq K$, we have $a_k \not\geq a_{k'} \not\geq a_k$.

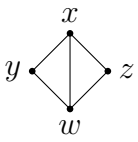


(a) A 4-crown. (b) A 6-crown. (c) An 8-crown.

Figure 1: Crowns.

transitive.[12] Neither is necessary:

Example 3. Consider $\mathcal{X}$ with the partial order $\gtrsim$ given by



That is: $x \gtrsim y \gtrsim w$ and $x \gtrsim z \gtrsim w$ (hence $x \gtrsim w$ by transitivity of $\gtrsim$), but y, z are $\gtrsim$-incomparable. $\gtrsim$ is not complete since y, z are $\gtrsim$-incomparable. Nor is $\gtrapprox$ transitive, since $y \gtrapprox x \gtrapprox z$ but y, z are $\gtrsim$-incomparable. But $\gtrsim$ is manifestly crown-free.

A small elaboration of the reasoning in Example 2 shows that the presence of a crown implies the existence of a pair of preferences with no minimum upper bound:

Lemma 2 (necessity of crown-freeness)**.** If every pair of preferences possesses a join, then $\gtrsim$ is crown-free.

Proof. We prove the contra-positive. Suppose that $\gtrsim$ contains a crown $(x_1, \ldots, x_K)$. Consider $P = \{\succeq_a, \succeq_b\} \subseteq \mathcal{P}$, where

$$x_K \succ_a x_1 \succ_a x_2 \succ_a \cdots \succ_a x_{K-2} \succ_a x_{K-1}$$
$$x_2 \succ_b x_3 \succ_b \cdots \succ_b x_{K-1} \succ_b x_K \succ_b x_1.$$

For $1 < k \leq K$ even, $x_{k-1} > x_k$ and $x_{k-1} \succ x_k$ for some $\succeq \in P$ (in particular, $\succeq_a$ for $k < K$ even, $\succeq_b$ for $k > 1$ even). Hence (x_{k-1}, x_k) is a strict P-chain, so by the characterisation theorem (p. 127), $x_{k-1} \succ^\star x_k$ for any minimum upper bound $\succeq^\star$ of P.

12. Given a binary relation $\geq$, comparability $\gtrapprox$ is defined by $a \gtrapprox b$ iff either $a \geq b$ or $b \geq a$.

Moreover, for $1 < k \leq K$ even, $x_{k+1} > x_k$, and $x_{k+1} \not\succeq x_k$ for all $\succeq \in P$. (This is apparent, separately, for $k < K$ even and for $k = K$.) Hence there is no P-chain from x_{k+1} to x_k, so by the characterisation theorem, $x_k \succ^\star x_{k+1}$ for any minimum upper bound of $\succeq^\star$ of P.

It follows that any minimum upper bound $\succeq^\star$ of P must satisfy

$$x_1 \succ^\star x_2 \succ^\star \cdots \succ^\star x_{K-1} \succ^\star x_K \succ^\star x_1.$$

Such a $\succeq^\star$ cannot be transitive, so cannot live in $\mathcal{P}$; hence P admits no minimum upper bound. ■

Application to politically correct discourse (continued)**.** Lemma 2 tells us that self-censorship may lead agents to express incoherent (i.e. intransitive) views. By way of illustration, consider the issues of economic justice for ethnic minorities (x), for young ethnic-majority men (y), for women (z) and for older ethnic-majority men (w). Political correctness $\gtrsim$ is plausibly as in Example 2: ethnic minorities and women are favoured over ethnic-majority men (both x and z $\gtrsim$-dominate both y and w), and the relative importance of ethnic minorities and women and of ethnic-majority men of different ages is indeterminate (x, z and y, w $\gtrsim$-incomparable). That is: (x, y, z, w) is a crown.

Agent 1 favours prioritising w over z, but since $z \gtrsim w$ and agent 2 holds the politically correct view $z \succ_2 w$, agent 1 expresses the view that $z \succ_1' w$. On all other pairs of issues, she expresses her true view; in particular, she favours w over x, x over y and y over z. Her expressed view $\succeq_1'$ therefore satisfies

$$w \succ_1' x \succ_1' y \succ_1' z \succ_1' w,$$

so is incoherent. Agent 2's post-conversation preference is incoherent for a similar reason.

2.2.2 Diamonds

Existence can fail even in the absence of crowns:

Example 3 (continued)**.** Let $P = \{\succeq_1, \succeq_2\} \subseteq \mathcal{P}$, where

$$y \succ_1 w \succ_1 z \succ_1 x \quad \text{and} \quad w \succ_2 z \succ_2 x \succ_2 y.$$

There is a strict P-chain from x to w (viz. (x, y, w)). We have $z \gtrsim w$ but no P-chain from z to w, and $x \gtrsim z$ but no P-chain from x to z. So by the characterisation theorem (p. 127), a minimum upper bound $\succeq^\star$ of P must satisfy

$$x \succ^\star w \succ^\star z \succ^\star x.$$

Since such a relation $\succeq^\star$ cannot be transitive, it follows that P admits no minimum upper bound.

The trouble is that $\gtrsim$ contains a *diamond:*

Definition 4. Let $\geq$ be a partial order on a set $\mathcal{A}$. A *diamond* is four elements (a, b, c, d) of $\mathcal{A}$ such that $a \geq b \geq d$ and $a \geq c \geq d$, but b, c are $\geq$-incomparable. $\geq$ is *diamond-free* iff it contains no diamonds.

Example 3 (continued)**.** (x, y, z, w) is a diamond: $x \gtrsim y \gtrsim w$ and $x \gtrsim z \gtrsim w$, but y, z are $\gtrsim$-incomparable. $\gtrsim$ therefore fails to be diamond-free.

Like crown-freeness, diamond-freeness rules out a specific form of incompleteness, and is implied by strong forms of 'limited completeness' such as completeness of $\geq$ or transitivity of the comparability relation $\gtrless$. Neither of these conditions is necessary for diamond-freeness, nor is crown-freeness:

Example 2 (continued)**.** We saw that there is a crown. But by inspection, there are no diamonds. Since $\gtrsim$ is not complete and $\gtrless$ is intransitive, this also shows that diamond-freeness is weaker than these sufficient conditions.

Lemma 3 (necessity of diamond-freeness)**.** If every pair of preferences possesses a join, then $\gtrsim$ is diamond-free.

The proof is almost exactly Example 3, so we omit it.

2.2.3 The existence theorem

We have seen in the last two sections that crown-freeness and diamond-freeness are individually necessary for every set of preferences to possess a minimum upper bound. The existence theorem asserts that these conditions are also jointly sufficient:

Existence theorem. The following are equivalent:

(1) Every *set* of preferences has a minimum upper bound.

(2) Every *pair* of preferences has a minimum upper bound.

(3) $\succsim$ is crown- and diamond-free.

Remark 1. This is a 'forbidden subposet' theorem: it characterises the class of posets $(\mathcal{X}, \succsim)$ on which all sets of preferences admit minimum upper bounds as precisely those that contain no crowns or diamonds as subposets.[13]

By way of illustration, crown- and diamond-freeness fails in Examples 2 and 3, but is satisfied in Example 1 (p. 125). More generally, it holds whenever there are three or fewer alternatives, and fails for any lattice $(\mathcal{X}, \succsim)$ that is not a chain.[14] To better understand the existence theorem, we provide a graphical interpretation of crown- and diamond-freeness in appendix I.

In appendix F, we show further that these properties are equivalent to every set (or two-element set) of preferences possessing a maximum lower bound. By analogy with complete lattices, call $(\mathcal{P}, S)$ a *complete pre-lattice* (a *pre-lattice*) iff every set $P \subseteq \mathcal{P}$ (every two-element set $P \subseteq \mathcal{P}$) possesses a minimum upper bound and maximum lower bound.[15] The existence theorem implies that $(\mathcal{P}, S)$ is a complete pre-lattice iff it is a pre-lattice iff $\succsim$ contains no crowns or diamonds.

13. A *subposet* of $(\mathcal{X}, \succsim)$ is $(\mathcal{Y}, \succsim_{\mathcal{Y}})$ where $\mathcal{Y} \subseteq \mathcal{X}$ and $\succsim_{\mathcal{Y}}$ is the restriction of $\succsim$ to $\mathcal{Y}$.

14. If $(\mathcal{X}, \succsim)$ is a lattice but not a chain, then there are $\succsim$-incomparable $x, y \in \mathcal{X}$, in which case $(x \wedge y, x, y, x \vee y)$ is a diamond.

15. $(\mathcal{P}, S)$ is not a complete lattice because S is not anti-symmetric in general, so that S fails to be a partial order. S is a pre-order (reflexive and transitive), though; hence the term 'complete pre-lattice'. The failure of anti-symmetry means that minimum upper bounds and maximum lower bounds need not be unique; we study this issue in §2.3.

As for the proof, it is immediate that (1) implies (2), and we have already established in Lemmata 2 and 3 that (2) implies (3). Proving that (3) implies (1) is more difficult; we do this in appendix D. The idea is as follows. Let $\succeq^\circ$ be the minimal binary relation (in general incomplete) that satisfies properties ($\star$)–($\star\star$) in the characterisation theorem (p. 127). We show first that absent diamonds in $\gtrsim$, $\succeq^\circ$ must be 'weakly transitive'. We then show that when there are no crowns in $\gtrsim$, weak transitivity of $\succeq^\circ$ implies that it satisfies a stronger transitivity-type property called *Suzumura-consistency*. This permits us to invoke an extension theorem due to Suzumura (1976) (stated in appendix B) to conclude that $\succeq^\circ$ can be extended to a complete and transitive relation (i.e. a preference). This preference is a minimum upper bound by the characterisation theorem.

Application to politically correct discourse (continued)**.** We saw that in Example 2, where $\gtrsim$ contains a crown, the agents' expressed views are incoherent. The same is true in Example 3, where $\gtrsim$ contains a diamond. In these examples, the need to self-censor causes agents to express incoherent views despite the fact that their true views are coherent.

The existence theorem tells us that that the expressed views in a conversation between *any* two agents $P = \{\succeq_1, \succeq_2\}$ are coherent if, and only if, $\gtrsim$ has neither crowns nor diamonds. In other words, the societies in which self-censorship leaves room for coherent expressions of opinion are precisely those in which the norms $\gtrsim$ of political correctness are crown- and diamond-free.

Crowns and diamonds can plausibly occur in political correctness $\gtrsim$. We already gave an example of four issues that are plausibly a crown. As for diamonds, the following issues are a plausible example: economic justice for ethnic-minority women (x), ethnic-minority men (y), ethnic-majority women (z) and ethnic-majority men (w). Self-censorship may therefore plausibly lead to incoherent expressed views in some conversations.

2.3 Uniqueness of minimum upper bounds

The existence theorem tells us when minimum upper bounds are guaranteed to exist. They are not unique, in general:

Example 4. Consider $\mathcal{X} = \{x, y\}$ with the empty partial order $\gtrsim$, so that x, y are $\gtrsim$-incomparable. Let $P = \{\succeq_1, \succeq_2\} \subseteq \mathcal{P}$, where $x \succ_1 y$ and $y \succ_2 x$.

Since all alternatives are $\gtrsim$-incomparable, there are no P-chains. Conditions ($\star$)–($\star\star$) in the characterisation theorem (p. 127) are therefore (vacuously) satisfied by any preference. So by the characterisation theorem, every preference is a minimum upper bound of P.

The message of Example 4 is that preferences that disagree only on $\gtrsim$-incomparable pairs of alternatives S-dominate each other, leading to multiplicity of minimum upper bounds for some sets of preferences. The solution is to rule out $\gtrsim$-incomparability:

Uniqueness proposition. The following are equivalent:

(1) Every set of preferences has *at most* one minimum upper bound.

(2) Every set of preferences has *exactly* one minimum upper bound.

(3) $\gtrsim$ is complete.

A strengthening of this result involving maximum lower bounds is given in appendix F. The stronger result implies that $(\mathcal{P}, S)$ is a complete lattice iff it is a lattice iff $\gtrsim$ is complete.

Proof that (3) implies (2). Suppose that $\gtrsim$ is complete, and fix a set $P \subseteq \mathcal{P}$. Since completeness implies crown- and diamond-freeness, P has at least one minimum upper bound by the existence theorem (p. 134).

To show uniqueness, let $\succeq', \succeq'' \in \mathcal{P}$ be minimum upper bounds of $P \subseteq \mathcal{P}$. Then by the characterisation theorem (p. 127), $\succeq'$ and $\succeq''$ must agree on all $\gtrsim$-comparable pairs of alternatives. Since $\gtrsim$ is complete, it follows that $\succeq'$ and $\succeq''$ agree on all pairs of alternatives, i.e. that they are identical. ■

It is immediate that (2) implies (1). The final implication requires the following lemma, proved in appendix E.

Lemma 4. Let $x, y \in \mathcal{X}$ be $\gtrsim$-incomparable. Then any set $P \subseteq \mathcal{P}$ has upper bounds $\succeq', \succeq'' \in \mathcal{P}$ such that $x \succ' y$ and $y \succ'' x$.

Proof that (1) implies (3). We prove the contra-positive. Suppose that $\gtrsim$ is incomplete, so that some $x, y \in \mathcal{X}$ are $\gtrsim$-incomparable; we seek a $P \subseteq \mathcal{P}$ with multiple upper bounds. By Lemma 4, such a P can be found. (Indeed, *every* $P \subseteq \mathcal{P}$ has multiple upper bounds.) ■

Application to politically correct discourse (continued)**.** Agents need not express consensus: they may well express views that are distinct minimum upper bounds of their true views $P = \{\succeq_1, \succeq_2\}$. In Example 4, political correctness makes no demands, so agents express their differing true views.

The uniqueness proposition tells us that *any* two agents $P = \{\succeq_1, \succeq_2\}$ will express consensus if, and only if, $\gtrsim$ is complete. In other words, self-censorship fully extinguishes expressions of differing views only when political correctness fully specifies how social issues ought to be ranked.

We have seen examples of issues for which political correctness $\gtrsim$ is plausibly incomplete. For some issues, though, completeness may obtain: an example is economic justice for women (x), for men (y) and for dogs (z), where $x \gtrsim y \gtrsim z$. Conversations on this issue will lead any pair of agents to express concordant views, even if they actually disagree.

3 Application to monotone comparative statics

In this section, we extend the theory of monotone comparative statics[16] in two directions. First, in §3.2, we consider comparative statics for the *consensus:* under what circumstances does the set of alternatives that every individual in a group $P \subseteq \mathcal{P}$ considers optimal increase? Second, in §3.3, we consider *misspecification-robust* comparative statics: under what circumstances can an analyst who knows only that the agent's preference lies in a set $P \subseteq \mathcal{P}$ conclude that optimal alternatives will increase?

Our Proposition 1 (§3.2) establishes that when P increases in the strong set order, so does the consensus. Our Proposition 2 (§3.3) similarly shows that when P increases in the *alternative* set order (a natural weakening of the strong set order), so does the set of possibly-optimal alternatives. Both proofs hinge on the existence theorem.

16. See Topkis (1978), Milgrom and Shannon (1994) and LiCalzi and Veinott (1992).

3.1 The canonical theory

We begin with a brief recap of the canonical theory of monotone comparative statics. An agent chooses an alternative x from a set $\mathcal{X} \subseteq \mathbf{R}$ ordered by the usual inequality $\geq$. The agent chooses optimally with respect to her preference $\succeq \in \mathcal{P}$.[17] Denote by $X(\succeq)$ the (possibly empty) set of optimal alternatives for preference $\succeq \in \mathcal{P}$:

$$X(\succeq) \coloneqq \{x \in \mathcal{X} : x \succeq y \text{ for every } y \in \mathcal{X}\}.$$

Definition 5. Consider a lattice $(\mathcal{A}, \geq)$ and two subsets $A, B \subseteq \mathcal{A}$. A dominates B in the *($\geq$-induced) strong set order* iff for any $a \in A$ and $b \in B$, we have $a \vee b \in A$ and $a \wedge b \in B$.

Note well that A dominates B in the strong set order whenever either set is empty.

MCS theorem.[18] Let $\succeq, \succeq' \in \mathcal{P}$ be preferences. If $\succeq' S \succeq$, then $X(\succeq')$ dominates $X(\succeq)$ in the ($\geq$-induced) strong set order.

In other words, when the agent's preference increases in the sense of single-crossing dominance S, the set of optimal alternatives increases in the sense of the strong set order.

Remark 2. There is a variant of the MCS theorem that applies whenever $(\mathcal{X}, \gtrsim)$ is a lattice. This added generality is not useful for our purposes because in order to apply the existence theorem, we shall require that $\gtrsim$ be crown- and diamond-free, and the only diamond-free lattices are chains (totally ordered sets).[19] It is therefore essentially without loss of generality to assume that $\mathcal{X}$ is a subset of $\mathbf{R}$, ordered by inequality $\geq$.[20]

3.2 Comparative statics for the consensus

Consider a problem in which there is a group of agents, each with a preference $\succeq \in \mathcal{P}$. Write $P \subseteq \mathcal{P}$ for the set of preferences in the group. The *consensus* $C(P)$ is the set of

17. This setting is more frequently studied using a utility representation rather than the preference itself, but this makes no material difference.

18. Milgrom and Shannon (1994) and LiCalzi and Veinott (1992).

19. See section 2.2.3 (p. 134).

20. It is only 'essentially' without loss because there are chains $(\mathcal{X}, \gtrsim)$ that cannot be embedded in $(\mathbf{R}, \geq)$ because the cardinality of $\mathcal{X}$ exceeds that of $\mathbf{R}$.

alternatives that *every* individual in the group finds optimal:

$$C(P) \coloneqq \bigcap_{\succeq \in P} X(\succeq).$$

Since $\geq$ is complete, $(\mathcal{P}, S)$ is a lattice by the uniqueness proposition (p. 136). We may therefore use the (S-induced) strong set order to compare sets of preferences.

Proposition 1 (consensus comparative statics)**.** Let $\mathcal{X}$ be a subset of $\mathbf{R}$ ordered by inequality $\geq$, and let $P, P' \subseteq \mathcal{P}$ be sets of preferences. If P' dominates P in the (S-induced) strong set order, then $C(P')$ dominates $C(P)$ in the ($\geq$-induced) strong set order.

In other words, when agents' preferences shift up in the sense of the strong set order, so does the consensus. Note that it may be that either $C(P)$ or $C(P')$ is empty, in which case the conclusion holds automatically.

Proof. Fix $P, P' \subseteq \mathcal{P}$ such that P' dominates P in the S-induced strong set order. The conclusion is immediate if either $C(P)$ or $C(P')$ is empty, so suppose not. Take $x \in C(P)$ and $x' \in C(P')$; we must show that $x \wedge x'$ lies in $C(P)$ and that $x \vee x'$ lies in $C(P')$. We will prove the former; the proof of the latter is similar.

Take any $\succeq \in P$ and $\succeq' \in P'$. Since the order $\geq$ on $\mathcal{X}$ is complete, the set $\{\succeq, \succeq'\}$ possesses a join $\succeq^\star$ by the uniqueness proposition (p. 136). Since P' dominates P in the S-induced strong set order, the join $\succeq^\star$ lies in P'. Because $\succeq^\star \; S \; \succeq'$, the MCS theorem implies that $X(\succeq^\star)$ dominates $X(\succeq)$ in the $\geq$-induced strong set order. Since $x \in C(P) \subseteq X(\succeq)$ and $x' \in C(P') \subseteq X(\succeq^\star)$, it follows that $x \wedge x' \in X(\succeq)$. Since $\succeq \in P$ was arbitrary, this shows that $x \wedge x' \in C(P)$. ■

Proposition 1 can be used to study comparative statics for social choice. A *social choice function (SCF)* is a map $\phi : 2^{\mathcal{P}} \setminus \{\varnothing\} \to \mathcal{X}$ that picks an alternative for each non-empty set of preferences.[21] A SCF ϕ is *monotone* iff $\phi(P') \geq \phi(P)$ whenever P' dominates P in

21. It is more typical to have a SCF map *profiles* of preferences into alternatives, but the distinction is immaterial for our purposes.

the (S-induced) strong set order, and *respects unanimity* iff $\phi(P) \in C(P)$ whenever the latter is non-empty.

Corollary 1. Let $\mathcal{X} \subseteq \mathbf{R}$ be compact. Then there exists a monotone SCF that respects unanimity.

The (not entirely trivial) proof is in appendix G.

3.3 Misspecification-robust comparative statics

Suppose that an analyst wishes to obtain a comparative statics conclusion, but is unsure of the agent's preference. In particular, she knows only that the agent's preference initially lies in $P \subseteq \mathcal{P}$, and that it subsequently lies in $P' \subseteq \mathcal{P}$. Under what circumstances can she obtain a comparative statics conclusion that is robust to misspecification of the agent's preference?

Write

$$X(P) \coloneqq \bigcup_{\succeq \in P} X(\succeq)$$

for the set possibly-optimal alternatives, given that the agent has *some* preference in P. The analyst wishes to infer that when P increases in a suitable sense, $X(P)$ also increases in a suitable sense. The following variant of the strong set order turns out to be appropriate:

Definition 6. Consider a lattice $(\mathcal{A}, \geq)$ and two subsets $A, B \subseteq \mathcal{A}$. A dominates B in the *($\geq$-induced) alternative set order* iff for any $a \in A$ and $b \in B$, there is an $a' \in A$ such that $a' \geq a \vee b$ and a $b' \in B$ such that $a \wedge b \geq b'$.

The alternative set order is formally weaker than the strong set order, but intuitively captures the same notion of 'larger'.

Proposition 2 (misspecification-robust comparative statics)**.** Let $\mathcal{X}$ be a subset of $\mathbf{R}$ ordered by inequality $\geq$, and let $P, P' \subseteq \mathcal{P}$ be non-empty sets of preferences such that $X(\succeq)$ is non-empty for every $\succeq \in P \cup P'$. If P' dominates P in the (S-induced) alternative set order, then $X(P')$ dominates $X(P)$ in the ($\geq$-induced) alternative set order.

The assertion becomes false if 'alternative' is replaced with 'strong'. The hypothesis that the argmaxes $X(\succeq)$ are non-empty is unfortunately essential, as we show by example in appendix J. It can be ensured by standard sufficient conditions.[22]

Proof. Fix $P, P' \subseteq \mathcal{P}$ such that P' dominates P in the S-induced alternative set order. Take $x \in X(P)$ and $x' \in X(P')$. We will show that there is a $y' \geq x \vee x'$ that lies in $X(P')$ and a $y \leq x \wedge x'$ that lies in $X(P)$.

By hypothesis, there are preferences $\succeq \in P$ and $\succeq' \in P'$ such that $x \in X(\succeq)$ and $x' \in X(\succeq')$. Since the order $\geq$ on $\mathcal{X}$ is complete, these preferences possess a minimum upper bound $\succeq \vee \succeq'$ and a maximum lower bound $\succeq \wedge \succeq'$ by the uniqueness proposition (p. 136). Since P' dominates P in the S-induced alternative set order, there is a $\succeq^\star \in P'$ such that $\succeq^\star S (\succeq \vee \succeq')$ and a $\succeq_\star \in P$ such that $(\succeq \wedge \succeq') S \succeq_\star$.

Since

$$\succeq^\star S (\succeq \vee \succeq') S \succeq S (\succeq \wedge \succeq') S \succeq_\star \quad \text{and similarly} \quad \succeq^\star S \succeq' S \succeq_\star,^{23}$$

the MCS theorem yields

$$X(\succeq^\star) \sqsupseteq X(\succeq) \sqsupseteq X(\succeq_\star) \quad \text{and} \quad X(\succeq^\star) \sqsupseteq X(\succeq') \sqsupseteq X(\succeq_\star),$$

where $\sqsupseteq$ denotes the $\geq$-induced strong set order on subsets of $\mathcal{X}$. It follows that

$$X(\succeq^\star) \sqsupseteq X(\succeq) \vee X(\succeq') \quad \text{and} \quad X(\succeq) \wedge X(\succeq') \sqsupseteq X(\succeq_\star),$$

where

$$X(\succeq) \vee X(\succeq') \coloneqq \{y \vee y' : y \in X(\succeq) \text{ and } y' \in X(\succeq')\},$$

and $X(\succeq) \wedge X(\succeq')$ is defined analogously.

Since $X(\succeq^\star)$ is non-empty by hypothesis, there is a $x^\star \in X(\succeq^\star)$. Since $X(\succeq^\star) \sqsupseteq X(\succeq) \vee X(\succeq')$, $y' \coloneqq x^\star \vee (x \vee x')$ lies in $X(\succeq^\star)$, and hence in $X(P')$ since $\succeq^\star \in P'$. Clearly

22. For example, finiteness of $\mathcal{X}$, compactness of $\mathcal{X}$ plus a continuity hypothesis, or the purely order-theoretic sufficient conditions of Milgrom and Shannon (1994, Theorem A.4).

23. We are relying on the fact that S is transitive.

$y' \geq x \vee x'$, as required.

Similarly, we may choose $x_\star \in X(\succeq_\star)$, set $y \coloneqq x_\star \wedge (x \wedge x')$, and conclude that y lies in $X(\succeq_\star) \subseteq X(P)$, and that $x \wedge x' \geq y$. ■

4 Application to ambiguity- and risk-aversion

In this section, we apply our results to ambiguity-aversion,[24] and in particular to the comparison of preferences according to 'more ambiguity-averse than'. We introduce and study a general class of maxmin preferences, and characterise this class in terms of minimum upper bounds.

In §4.2, we introduce *(generalised) maxmin preferences,* which (greatly) generalise maxmin expected utility (Gilboa and Schmeidler 1989). In §4.3, we use our theorems to prove that maxmin preferences may be characterised as minimum upper bounds with respect to 'more ambiguity-averse than'. In §4.4, we adapt the argument to preferences over *lotteries.*

4.1 Environment

The setting is the Savage (1954) framework, with money prizes as consequences. There is a set Ω of *states of the world* and a set $\Pi \subseteq \mathbf{R}$ of monetary *prizes.*[25] Let $\mathcal{X}$ be an arbitrary set of *acts,* meaning functions $\Omega \to \Pi$,[26] with typical elements $X, Y \in \mathcal{X}$. Write $\mathcal{C} \subseteq \mathcal{X}$ for subset of *constant acts,* meaning those that always lead to the same prize. Typical constant acts are denoted by $C, D \in \mathcal{C}$. By a standard abuse of notation, we identify each constant act $C \in \mathcal{C}$ with the prize (element of $\Pi \subseteq \mathbf{R}$) at which it is constant. Assume for that $\mathcal{C}$ (viewed as a subset of $\mathbf{R}$) is closed.

Let $\mathcal{P}$ be the set of all preferences over $\mathcal{X}$. A preference is called more ambiguity-averse than another iff whenever the latter prefers a constant act, the former does, too:

24. See Ellsberg (1965), Schmeidler (1989), Gilboa and Schmeidler (1989), Klibanoff, Marinacci and Mukerji (2005), Maccheroni, Marinacci and Rustichini (2006) and Cerreia-Vioglio et al. (2011).
25. The analysis is easily extended to allow for objective lotteries over monetary prizes.
26. We do *not* restrict attention to acts with finite co-domain.

Definition 7. For two preferences $\succeq, \succeq' \in \mathcal{P}$, we say that $\succeq'$ is *more ambiguity-averse than* $\succeq$ iff for an act $X \in \mathcal{X}$ and a constant act $C \in \mathcal{C}$, $C \succeq (\succ) X$ implies $C \succeq' (\succ') X$.

This definition is standard—see Ghirardato and Marinacci (2002) and Epstein (1999).[27] It is inspired by Yaari's (1969) definition of 'more risk-averse than', which we will study in §4.4.

Say that a preference $\succeq \in \mathcal{P}$ is *monotone* iff it likes money: for any pair of constant acts $C, D \in \mathcal{C}$, we have $C \succeq D$ iff $C \geq D$. A *certainty equivalent* for $\succeq \in \mathcal{P}$ of an act $X \in \mathcal{X}$ is a constant act $C \in \mathcal{C}$ such that $X \succeq C \succeq X$. A preference is called *solvable* iff it has a certainty equivalent for every act.[28]

Let $\mathcal{P}' \subseteq \mathcal{P}$ denote the set of monotone and solvable preferences over $\mathcal{X}$. Write $c(\succeq, X)$ for the (by monotonicity, unique) certainty equivalent of an act $X \in \mathcal{X}$ according to a preference $\succeq \in \mathcal{P}'$.

A function $U : \mathcal{X} \to \mathbf{R}$ *ordinally represents* a preference $\succeq \in \mathcal{P}$ iff for any acts $X, Y \in \mathcal{X}$, we have $X \succeq Y$ iff $U(X) \geq U(Y)$. Given a preference $\succeq \in \mathcal{P}$, a strictly increasing (utility) function $u : \Pi \to \mathbf{R}$ and a probability measure (a belief) μ on Ω, we say that (u, μ) is a *subjective expected utility (SEU) representation* of $\succeq$ iff

$$X \mapsto \int_{\Omega} (u \circ X) \, \mathrm{d}\mu$$

ordinally represents $\succeq$.[29] $\succeq$ is called a *SEU preference* iff such a pair (u, μ) can be found.

A preference $\succeq \in \mathcal{P}$ is called *probabilistically sophisticated* iff it can be viewed as arising from a decision-maker with a belief $\mu \in \Delta(\Omega)$ about the state who cares about an act X solely through the (subjective) lottery over prizes $\mu \circ X^{-1}$ that it induces.[30] Formally, $\succeq \in \mathcal{P}$ is probabilistically sophisticated iff there exists a probability measure

27. Our analysis carries over to the case where, in the sprit of Epstein (1999), $\mathcal{C}$ is the set of all 'unambiguous' acts, including but not necessarily limited to constant acts.

28. In the sequel, it is natural (though not necessary) to assume that $\mathcal{C}$ contains *every* constant function $\Omega \to \Pi$.

29. We shall not be careful about what exactly is meant by the integral '$\int$', since this detail is irrelevant for our results.

30. The *pushforward measure* $\mu \circ X^{-1}$ is the measure on Π defined by $\left(\mu \circ X^{-1}\right)(A) \coloneqq \mu(\{\omega \in \Omega : X(\omega) \in A\})$ for each (measurable) $A \subseteq \Pi$.

$\mu \in \Delta(\Omega)$ and a preference $\succeq_L$ over monetary lotteries $\Delta(\Pi)$ such that

$$\text{for acts } X, Y \in \mathcal{X}, \quad X \succeq Y \quad \text{iff} \quad \mu \circ X^{-1} \succeq_L \mu \circ Y^{-1}.$$

Any SEU preference is probabilistically sophisticated, but the converse is false.

4.2 Generalised maxmin preferences

We are concerned with preferences over acts that admit a maxmin representation of a general sort:

Definition 8. A set $P \subseteq \mathcal{P}'$ of monotone and solvable preferences is a *(generalised) maxmin representation* of a preference $\succeq^\star \in \mathcal{P}$ iff

$$X \mapsto \inf_{\succeq \in P} c(\succeq, X)$$

ordinally represents $\succeq^\star$, where 'inf' is the ordinary infimum over real numbers.

Intuitively, such a decision-maker is unsure which preference $\succeq \in P$ to evaluate acts according to, so cautiously values acts at their smallest certainty equivalent among $\succeq \in P$. An alternative intepretation is that there is a group of agents with preferences P, and that collective decisions are made according to the 'Rawlsian' maxmin criterion.

A special case is a *maxmin expected utility representation:* this is a maxmin representation $P \subseteq \mathcal{P}'$ consisting entirely of SEU preferences that have the same utility function, i.e. there exists a $u : \Pi \to \mathbf{R}$ such that every preference $\succeq \in P$ has a SEU representation $(u, \mu_\succeq)$ for some $\mu_\succeq \in \Delta(\Omega)$.[31] Gilboa and Schmeidler (1989) introduced maxmin expected utility preferences, and gave axioms in the Anscombe–Aumann (1963)

31. To see explicitly that this is equivalent to the usual definition of maxmin expected utility, observe first that u must be strictly increasing since the preferences in P are monotone. On our definition, $\succeq^\star$ is ordinally represented by

$$X \mapsto \inf_{\succeq \in P} c(\succeq, X) = \inf_{\succeq \in P} u^{-1}\left(\int_\Omega (u \circ X) \mathrm{d}\mu_\succeq\right) = u^{-1}\left(\inf_{\succeq \in P} \int_\Omega (u \circ X) \mathrm{d}\mu_\succeq\right),$$

which is a strictly increasing transformation of $X \mapsto \inf_{\succeq \in P} \int_\Omega (u \circ X) \mathrm{d}\mu_\succeq$.

framework. Casadesus-Masanell, Klibanoff and Özdenören (2000) gave axioms in the Savage setting.

Another special case is *subjective cautious expected utility,* where P is a set of SEU preferences with the same belief $\mu \in \Delta(\Omega)$. In this case, the decision-maker is probabilistically sophisticated, but her implicit preference $\succeq_L$ over monetary lotteries has the 'cautious expected utility' form axiomatised by Cerreia-Vioglio, Dillenberger and Ortoleva (2015).[32]

The *entire* class of maxmin preference is too broad to put substantial restrictions on behaviour:

Observation 1 (maxmin axiomatisation)**.** For a preference $\succeq^\star \in \mathcal{P}$, the following are equivalent:

(1) $\succeq^\star$ is monotone and solvable.

(2) $\succeq^\star$ admits a maxmin representation.

Proof. Recall that by definition, a maxmin representation is a set $P \subseteq \mathcal{P}'$ of *monotone and solvable* preferences. It is easily verified that a preference that admits a maxmin representation must itself be monotone, and solvability follows since $\mathcal{C}$ (as a subset of $\mathbf{R}$) is closed. For the converse, suppose that $\succeq^\star$ is monotone and solvable; then $P \coloneqq \{\succeq^\star\}$ is a maxmin representation. ■

Interesting behaviour arises, however, when P is restricted beyond merely monotonicity and solvability, but not so severely as to yield maxmin expected utility or subjective cautious expected utility. For instance:

Example 5. Consider a decision-maker who has maxmin expected utility (MEU) preferences, but who faces 'background risk' in the form of uncertain wealth $W : \Omega \to \Pi$. Her

32. In this case, $\succeq^\star$ is ordinally represented by

$$X \mapsto \inf_{\succeq \in P} u_\succeq^{-1}\left(\int_\Omega (u_\succeq \circ X)\mathrm{d}\mu\right) = \inf_{\succeq \in P} u_\succeq^{-1}\left(\int_\Omega u_\succeq \mathrm{d}\left(\mu \circ X^{-1}\right)\right).$$

The implicit preference $\succeq_L$ over monetary lotteries $q \in \Delta(\Pi)$ therefore admits the cautious expected utility representation $q \mapsto \inf_{\succeq \in P} u_\succeq^{-1}\left(\int_\Omega u_\succeq \mathrm{d}q\right)$.

preference is ordinally represented by

$$X \mapsto \inf_{\succeq \in P} \int_{\Omega} (u \circ [X + W]) \mathrm{d}\mu_{\succeq},^{33}$$

where $P \subseteq \mathcal{P}'$ is a MEU representation, i.e. a set of SEU preferences that share the same (strictly increasing) risk attitude $u : \Pi \to \mathbf{R}$, but differ in their beliefs $\mu_{\succeq}$ on Ω. We call this MEU with background risk (MEU-BR).

The MEU-BR class is contained in the (generalised) maxmin class—simply observe that MEU-BR preferences satisfy monotonicity and solvability, and apply Observation 1. In fact, the containment is strict: MEU-BR preferences satisfy behavioural axioms in addition to monotonicity and solvability. For example, any MEU-BR preference $\succeq^{\star}$ clearly satisfies *state-wise monotonicity:* if $X, Y \in \mathcal{X}$ are such that $X(\omega) \geq Y(\omega)$ for every $\omega \in \Omega$, then $X \succeq^{\star} Y$.

On the other hand, MEU-BR strictly extends the MEU model, allowing for choice behaviour that the latter cannot rationalise. For example, while MEU preferences satisfy a form of independence called *ordered A-act-independence* (Casadesus-Masanell, Klibanoff and Özdenören 2000), MEU-BR preferences need not.[34]

4.3 Characterisation of maxmin preferences

Our theorems permit us to characterise maxmin preferences as precisely minimum upper bounds with respect to 'more ambiguity-averse than':

Proposition 3 (maxmin characterisation)**.** For a preference $\succeq^{\star} \in \mathcal{P}$ and a set $P \subseteq \mathcal{P}'$ of monotone and solvable preferences over acts, the following are equivalent:

(1) P is a maxmin representation of $\succeq^{\star}$.

33. Assume that Π is closed under addition, so that $X(\omega) + W(\omega)$ lies in Π for each $\omega \in \Omega$.

34. A counter-example is as follows (see Casadesus-Masanell, Klibanoff and Özdenören (2000) for definitions). Let $\Pi = \mathbf{R}$ and $\Omega = \{0, 1\}$. The MEU-BR representation consists of wealth $W(\omega) = \omega$ and two SEU representations (u, μ_0) and (u, μ_1), where u is the identity and μ_ω assigns probability 1 to state ω. Observe that there are two ordered non-null and non-universal events, $\{0\}$ and $\{1\}$; fix one, $A = \{\omega^{\star}\}$. Any acts X, Y, Z, X', Y' with the following properties violate part (i) of A-act-independence: $V(\omega^{\star}) > V(1 - \omega^{\star})$ for $V \in \{X, Y, Z\}$, and $X(\omega) > Y(\omega) \geq Z(\omega) + 1$ and $X'(\omega) = Y'(\omega) = Z(\omega) + \omega^{\star}$ for $\omega \in \{1, 2\}$.

(2) $\succeq^\star$ is a minimum upper bound of P with respect to 'more ambiguity-averse than'.

Proposition 3 furnishes an intuitive way to interpret maxmin preferences. To wit, a preference with maxmin representation P is precisely one that is more ambiguity-averse than any preference in P, but no more ambiguity-averse than that.

Proposition 3 also delivers comparative statics for ambiguity-aversion:

Corollary 2 (comparative statics)**.** Let $\succeq, \succeq' \in \mathcal{P}$ admit maxmin representations $P, P' \subseteq \mathcal{P}'$. If P' contains P, or if P' dominates P in the strong set order,[35] then $\succeq'$ is more ambiguity-averse than $\succeq$.

Another consequence of Proposition 3 is that every set $P \subseteq \mathcal{P}'$ of monotone and solvable preferences possesses exactly one minimum upper bound with respect to 'more ambiguity-averse than'. The existence of minimum upper bounds can actually be proved directly using the existence theorem, and does not require solvability. In particular, define a binary relation $\gtrsim$ on $\mathcal{X}$ by $X \gtrsim Y$ iff either (i) $X = Y$, or (ii) X is constant and Y is not, or (iii) X, Y are constant and $X > Y$. For any pair $\succeq, \succeq' \in \mathcal{P}$ of monotone preferences, $\succeq'$ is more ambiguity-averse than $\succeq$ iff $\succeq'$ dominates $\succeq$ in the single-crossing relation induced by $\gtrsim$. It is trivial that $\gtrsim$ is crown- and diamond-free; we may therefore invoke the existence theorem.

Proof of Proposition 3. Fix a set $P \subseteq \mathcal{P}'$ of monotone and solvable preferences. Let $\gtrsim$ be the crown- and diamond-free partial order defined in the previous paragraph.

Remark that P-chains have a simple form. There is a strict P-chain between any distinct pair of constant acts because every preference in P is monotone. There are no P-chains between non-constant acts, since there are no $\gtrsim$-monotone sequences between them. Finally, there is a (strict) P-chain from a constant act C to a non-constant act X iff $C \succeq (\succ)\ X$ for some $\succeq \in P$.

We need therefore only consider P-chains of length 2 from a constant act to a non-constant one. So fix a constant $C \in \mathcal{C}$ and a non-constant $X \in \mathcal{X} \setminus \mathcal{C}$.

35. The strong set order was defined in §3.1 (p. 138).

(1) implies (2): Suppose that P is a maxmin representation of $\succeq^\star \in \mathcal{P}$. We will show that $\succeq^\star$ is a minimum upper bound of P by establishing that it satisfies conditions $(\star)$–$(\star\star)$ in the characterisation theorem (p. 127).

Suppose that there is a (strict) P-chain from C to X, which we argued implies that $C \succeq (\succ)\ X$ for some $\succeq \in P$. Then $c(\succeq, C) \geq (>)\ c(\succeq, X)$ for some $\succeq \in P$, and since every preference agrees on the certainty equivalent of a constant act, it follows that

$$\inf_{\succeq \in P} c(\succeq, C) \geq (>) \inf_{\succeq \in P} c(\succeq, X).$$

Thus $C \succeq^\star (\succ^\star)\ X$ since P is a maxmin representation of $\succeq^\star$.

Suppose that there is no (strict) P-chain from C to X: $C \not\succeq (\not\succ)\ X$ for every $\succeq \in P$. Then $c(\succeq, C) < (\leq)\ c(\succeq, X)$ for every $\succeq \in P$, so

$$\inf_{\succeq \in P} c(\succeq, C) < (\leq) \inf_{\succeq \in P} c(\succeq, X).$$

Since P is a maxmin representation of $\succeq^\star$, it follows that $C \not\succeq^\star (\not\succ^\star)\ X$.

(2) implies (1): Let $\succeq^\star \in \mathcal{P}$ be a minimum upper bound of P. We wish to show that P is a maxmin representation of $\succeq^\star$.

Suppose that $C \succeq^\star (\succ^\star)\ X$. Then by the characterisation theorem, there is a (strict) P-chain from C to X, which we argued means that $C \succeq (\succ)\ X$ for some $\succeq \in P$, or equivalently $c(\succeq, C) \geq (>)\ c(\succeq, X)$ for some $\succeq \in P$. Since all preferences agree on the certainty equivalent of a constant act, it follows that

$$\inf_{\succeq \in P} c(\succeq, C) \geq (>) \inf_{\succeq \in P} c(\succeq, X).$$

Suppose that $X \succ^\star (\succeq^\star)\ C$, or equivalently $C \not\succeq^\star (\not\succ^\star)\ X$. Then by the characterisation theorem, there is no (strict) P-chain from C to X. Therefore $c(\succeq, C) < (\leq)\ c(\succeq, X)$ for every $\succeq \in P$, so

$$\inf_{\succeq \in P} c(\succeq, C) < (\leq) \inf_{\succeq \in P} c(\succeq, X). \qquad \blacksquare$$

Savage (§4.1–4.3)	von Neumann–Morgenstern (§4.4)
act	lottery
constant act	degenerate lottery
certainty equivalent	certainty equivalent
more ambiguity-averse than	more risk-averse than
maxmin representation	cautious representation

Table 1: Relabelling of terms for §4.4.

4.4 Risk-aversion and caution

In this section, we carry our results over to preferences over monetary *lotteries* in the von Neumann–Morgenstern (1947) framework. In brief, we relabel terms according to Table 1, and thence obtain a characterisation of *cautious* preferences as precisely minimum upper bounds with respect to 'more risk-averse than'.

There is a set $\Pi \subseteq \mathbf{R}$ of monetary prizes and a set $\mathcal{X}$ of *lotteries* (probability distributions) over prizes. A *degenerate lottery*, $C \in \mathcal{C} \subseteq \mathcal{X}$, is one that pays out some prize $x \in \Pi$ with probability one. As is standard, we identify a degenerate lottery with the prize it pays. We write $\mathcal{P}$ for the set of all preferences over $\mathcal{X}$. A function $u : \Pi \to \mathbf{R}$ is called an *expected utility (EU) representation* of $\succeq \in \mathcal{P}$ iff $X \mapsto \int_\Pi u \mathrm{d}X$ ordinally represents $\succeq$.

Comparative risk-aversion is standardly defined as follows (Yaari 1969):

Definition 9. For two preferences $\succeq, \succeq' \in \mathcal{P}$, we say that $\succeq'$ is *more risk-averse than* $\succeq$ iff for a lottery $X \in \mathcal{X}$ and a degenerate lottery $C \in \mathcal{C}$, $C \succeq (\succ) X$ implies $C \succeq' (\succ') X$.

A preference $\succeq \in \mathcal{P}$ is *monotone* iff for any pair of degenerate lotteries $C, D \in \mathcal{C}$, we have $C \succeq D$ iff $C \geq D$. A *certainty equivalent* for $\succeq \in \mathcal{P}$ of a lottery $X \in \mathcal{X}$ is a degenerate lottery $C \in \mathcal{C}$ such that $X \succeq C \succeq X$. Let $\mathcal{P}' \subseteq \mathcal{P}$ be all monotone and solvable preferences, and write $c(\succeq, X)$ for their certainty equivalents.

Definition 10. A set $P \subseteq \mathcal{P}'$ of monotone and solvable preferences is a *cautious repres-*

entation of a preference $\succeq^\star \in \mathcal{P}$ iff

$$X \mapsto \inf_{\succeq \in P} c(\succeq, X)$$

ordinally represents $\succeq^\star$, where 'inf' is the ordinary infimum over real numbers.

The special case in which $P \subseteq \mathcal{P}'$ contains only EU preferences is called *(objective) cautious expected utility;* such preferences have been axiomatised by Cerreia-Vioglio, Dillenberger and Ortoleva (2015). Cautious representations are much more general than this, however: all (and only) monotone and solvable preferences admit a cautious representation.

Corollary 3 (caution characterisation)**.** For a preference $\succeq^\star \in \mathcal{P}$ and a set $P \subseteq \mathcal{P}'$ of monotone and solvable preferences over lotteries, the following are equivalent:

(1) P is a cautious representation of $\succeq^\star$.

(2) $\succeq^\star$ is a minimum upper bound of P with respect to 'more risk-averse than'.

Corollary 3 implies comparative statics along the lines of Corollary 2, as well as the existence and uniqueness of minimum upper bounds of sets of monotone and solvable preferences. Existence does not require solvability.

5 Application to social choice

In this section, we apply our theorems to social choice in the presence of normative constraints that require unanimity in order for certain alternatives to be ranked above others. We ask what normative constraints are compatible with a suitable Pareto property.

Our (im)possibility theorem (§5.3) provides the necessary and sufficient condition on a normative constraint for there to exist a social welfare function that satisfies it as well as the Pareto criterion. Since the condition is strong, this can be viewed as an impossibility result in the spirit of Sen's (1970) 'impossibility of a Paretian liberal'; we discuss this relationship in §5.4. But our result also has an encouraging side: we exhibit some natural normative constraints which are compatible with the Pareto property.

5.1 Environment

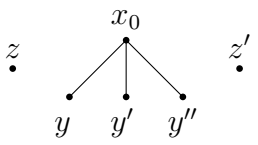
There is a society composed of n individuals, each with a preference over a set $\mathcal{X}$ of social alternatives. For certain pairs $x, y \in \mathcal{X}$, it is acceptable for society to rank y (strictly) above x only if all members of society consent; we write $x \gtrsim y$ when this is the case. We assume that $\gtrsim$ is a partial order (in particular, transitive).

For example, it may be that $x \gtrsim y$ if alternative involves more surveillance or social control than does x. Alternatively, consider a community deciding what amenities to provide at public expense. If alternative x differs from y in adding a luxury such as a swimming pool or gym, then it may be normatively acceptable for the community to favour y over x only if every individual consents, i.e. $x \gtrsim y$. Or it could be that $\gtrsim$ reflects the status quo:

Example 6. Suppose that there are old alternatives $\mathcal{Y}$ and new alternatives $\mathcal{Z}$, so that $\mathcal{X} = \mathcal{Y} \cup \mathcal{Z}$. In the past, when only alternatives $\mathcal{Y}$ were available, option $x_0 \in \mathcal{Y}$ was chosen. It is acceptable for society to depart from this status quo by ranking some $y \in \mathcal{Y}$ above x_0 only with the consent of all individuals: in other words, $x_0 \gtrsim y$ for every other $y \in \mathcal{Y}$.

Precedent does not constrain how previously-unchosen alternatives $y, y' \in \mathcal{Y} \setminus \{x_0\}$ are ranked, nor does it have anything to say about new alternatives $z \in \mathcal{Z}$. Formally, any alternatives $y, y' \in \mathcal{Y} \setminus \{x_0\}$ are $\gtrsim$-incomparable, and each $z \in \mathcal{Z}$ is $\gtrsim$-incomparable to every other alternative.

This fully specifies the partial order $\gtrsim$. It looks like this:



A *social welfare function (SWF)* is a map $f : \mathcal{P}^n \to \mathcal{P}$ that produces a social preference for any profile $\pi \in \mathcal{P}^n$ of individual preferences. This specification encodes the assumptions of *transitivity* (social preference must be transitive) and *unrestricted domain* (any profile of individual preferences is possible).

Definition 11. A SWF f *respects justified objections* iff for any profile $\pi = (\succeq_1, \ldots, \succeq_n)$ and alternatives $x \gtrsim y$, $x\ f(\pi)\ y$ (and not $y\ f(\pi)\ x$) whenever $x \succeq_i (\succ_i)\ y$ for some individual i.

Observe that a SWF f respects justified objections precisely if $f(\pi)$ is an upper bound of π for every profile π of preferences.[36] By the UB characterisation lemma (p. 126), this means precisely that for $x \gtrsim y$, $x\ f(\pi)\ y$ (and not $y\ f(\pi)\ x$) whenever there is a (strict) π-chain from x to y.

A SWF f *respects unanimity* (or 'is Pareto') iff for any profile $\pi = (\succeq_1, \ldots, \succeq_n)$, whenever $x \succeq_i (\succ_i)\ y$ for every individual i, we have $x\ f(\pi)\ y$ (and not $y\ f(\pi)\ x$). Unsurprisingly, respect for unanimity typically clashes with respect for justified objections:

Example 7. Let $\mathcal{X} = \{x, y, z\}$ with $x > y > z$, and let $\pi = (\succeq_1, \succeq_2)$, where $z \succ_1 x \succ_1 y$ and $y \succ_2 z \succ_2 x$. If f is a SWF that respects justified objections, then $f(\pi)$ must rank x above y since $x \gtrsim y$ and $x \succ_1 y$. $f(\pi)$ must similarly rank y above z, and thus rank x above z by transitivity. But then f fails to respect unanimity, since z is unanimously strictly preferred to x.

Unconditional respect for unanimity is therefore too demanding in the presence of normative constraints. It is more appropriate to ask that unanimity be respected only when this does not obviously conflict with respect for justified objections:

Definition 12. A SWF f *conditionally respects unanimity* iff for any profile $\pi = (\succeq_1, \ldots, \succeq_n)$, whenever $x \succeq_i (\succ_i)\ y$ for every individual i and there is no strict π-chain (no π-chain) from y to x, we have $x\ f(\pi)\ y$ (and not $y\ f(\pi)\ x$).

Call a SWF *acceptable* iff it respects justified objections and conditionally respects unanimity.

5.2 Acceptable SWFs need not exist

Since conditional respect for unanimity is designed to be compatible with respect for justified objections, it may appear that acceptable SWFs are guaranteed to exist. We will

36. We abuse notation throughout by identifying the profile $\pi = (\succeq_1, \ldots, \succeq_n)$ with the set $P \coloneqq \bigcup_{i=1}^{n} \{\succeq_i\}$ of preferences in it.

now see that this intuition is incorrect.

Lemma 5 (only joins are acceptable)**.** If f is acceptable, then $f(\pi)$ is a minimum upper bound of π for every profile $\pi \in \mathcal{P}^n$.

It follows by the existence theorem (p. 134) that a *necessary* condition for the existence of an acceptable SWF is that $\gtrsim$ be crown- and diamond-free. For instance, the preferences in Example 2 (p. 129) and in Example 3 (p. 132) cannot be aggregated in an acceptable way.

Proof. Suppose that f is acceptable, and fix a profile $\pi \in \mathcal{P}^n$. Respect for justified objections requires that $f(\pi)$ be an upper bound of π.

To see that conditional respect for unanimity requires minimumhood, suppose toward a contradiction that there is another upper bound $\succeq'$ of π such that $\succeq' \not\mathcal{S}\ f(\pi)$. Then there are alternatives $x, y \in \mathcal{X}$ with $x \gtrsim y$ such that y is (strictly) unanimously preferred to x and $y \succeq' (\succ') x$, but $x\ f(\pi)\ y$ (and not $y\ f(\pi)\ x$). Since $\succeq'$ is an upper bound of π, there cannot be a (strict) π-chain from x to y. Hence $f(\pi)$ fails to conditionally respect unanimity, so is unacceptable—a contradiction. ■

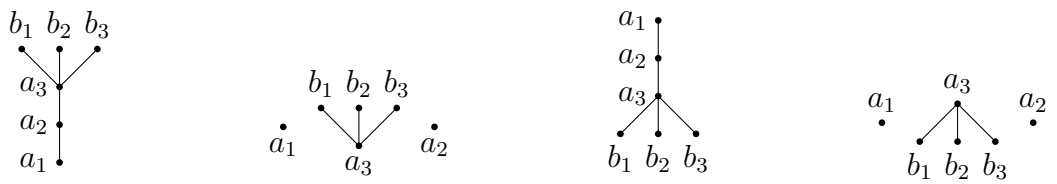
Crown- and diamond-freeness of $\gtrsim$ is not *sufficient* for acceptable preference aggregation, however:

Example 1 (continued)**.** Recall the details from p. 125. $\gtrsim$ is crown- and diamond-free by inspection.

Let there be two individuals $n = 2$, and let f be a SWF; we will show that f cannot be acceptable. Consider $\pi = (\succeq_1, \succeq_2)$, where $\succeq_1$ and $\succeq_2$ are as given previously (p. 125). Write $\succeq^\star \coloneqq f(\pi)$ for the social preference according to f.

By Lemma 5, a necessary condition for f to be acceptable is that $\succeq^\star$ be a minimum upper bound of π. We saw (p. 127) that this requires precisely that

$$w \succ^\star x \succ^\star y \succ^\star z.$$



(a) Up-fork. (b) Shattered fork. (c) Down-fork. (d) Shattered fork.

Figure 2: Transitive reductions of some forks and shattered forks.

But then f fails to conditionally respect unanimity: z is unanimously strictly preferred to w, and there is no P-chain from w to z, and yet $f(\pi) = \succeq^\star$ ranks w above z.

5.3 (Im)possibility theorem

We have shown that crown- and diamond-freeness is necessary, but not sufficient, for the existence of an acceptable SWF. It can be shown that completeness is sufficient but not necessary. Our (im)possibility theorem provides the necessary and sufficient condition: $(\mathcal{X}, \gtrsim)$ must be either a *fork* or a *shattered fork.*

Definition 13. Let $\geq$ be a binary relation on a set $\mathcal{A}$. An element $a \in \mathcal{A}$ is *isolated* iff it is $\geq$-incomparable to every other element of $\mathcal{A}$. A set $A \subseteq \mathcal{A}$ is an *upward (downard) fork head* iff it has a minimum (maximum) element $a \in A$ and $A \setminus \{a\}$ is an anti-chain, and simply a *fork head* iff it is either an upward or a downward fork head.

The poset $(\mathcal{A}, \geq)$ is an *up-fork (down-fork)* iff $\mathcal{A}$ can be partitioned as $\mathcal{A} = A \cup \{a\} \cup B$ where $A \cup \{a\}$ is an upward (downward) fork head with minimum (maximum) element a and $\{a\} \cup B$ is a chain with maximum (minimum) element a, and simply a *fork* iff it is an up-fork or a down-fork.

The poset $(\mathcal{A}, \geq)$ is a *shattered fork* iff $\mathcal{A}$ can be partitioned as $\mathcal{A} = A \cup B$, where A is a fork head and each $b \in B$ is isolated.

This definition is illustrated in Figure 2. Observe that $(\mathcal{X}, \gtrsim)$ being a fork or shattered fork implies that $\gtrsim$ is crown- and diamond-free, and is implied by completeness of $\gtrsim$. Since the converse implications do not hold, this property is strictly in-between the conditions for existence and for uniqueness of minimum upper bounds.

(Im)possibility theorem. The following are equivalent:

(1) There exists an acceptable SWF.

(2) $(\mathcal{X}, \gtrsim)$ is either a fork or a shattered fork.

The proof is in appendix H. There, we also provide a 'forbidden subposet' variant of the (im)possibility theorem according to which the existence of an acceptable SWF is equivalent to $\gtrsim$ being free of various subposets (including, but not limited to, crowns and diamonds). Note that as a special case, there exists an acceptable SWF whenever there are three or fewer alternatives.

The (im)possibility theorem bears both good news and bad. On the one hand, acceptable aggregation of preferences is impossible whenever $(\mathcal{X}, \gtrsim)$ is anything other than a fork or shattered fork, a stringent condition. On the other hand, if $(\mathcal{X}, \gtrsim)$ *is* a fork or shattered fork, then acceptable aggregation is possible.

It is easy to come up with examples in which the normative constraint is not a fork or shattered fork. In Examples 2 and 3, $\gtrsim$ fails to be crown- and diamond-free, so $(\mathcal{X}, \gtrsim)$ is certainly not a fork or shattered fork. Further:

Example 1 (continued)**.** Although $\gtrsim$ is crown- and diamond-free, $(\mathcal{X}, \gtrsim)$ is neither a fork nor a shattered fork by inspection. We saw that no acceptable SWF exists.

But equally, there are environments in which the fork-or-shattered-fork condition plausibly holds:

Example 6 (continued)**.** By inspection, the normative constraint $\gtrsim$ is a shattered (down-)fork. It follows by the (im)possibility theorem that there is an acceptable SWF.

Consider a variant in which there is a status quo ranking of the old alternatives $\mathcal{Y}$, and that the normative constraint requires unanimous consent to overturn any part of the status quo ranking. That is, $\mathcal{Y}$ is a chain in $\gtrsim$. Suppose further that unanimous consent is needed to rank a new alternative $z \in \mathcal{Z}$ above any old alternative $y \in \mathcal{Y}$, so that $y \gtrsim z$. (This is a very conservative society!) Then $(\mathcal{X}, \gtrsim)$ is a (down-)fork, so an acceptable SWF exists by the (im)possibility theorem.

5.4 Related literature

Our (im)possibility result can be thought of as complementary to Sen's (1970) 'impossibility of a Paretian liberal'. Sen defines 'minimal liberalism' as follows: there exist two individuals $i \neq j$ and four distinct alternatives x_i, y_i, x_j, y_j such that individual i (j) dictates the social preference on x_i, y_i (x_j, y_j). The idea is that x_i and y_i differ only in respects relevant only to individual i (such as the colour of i's wallpaper), and that such personal decisions should be devolved to the individual in a liberal society. Sen's theorem states that, when there are $n \geq 2$ individuals and $|\mathcal{X}| \geq 4$ alternatives, there exists no SWF that respects unanimity and minimal liberalism.

Sen's notion of liberalism is sensible when social alternatives specify *personal* matters (such as wallpaper colour). It is inappropriate for environments, such as surveillance, in which social alternatives are not such complete descriptions of the state of society.[37] In such contexts, our notion of respect for justified objections is a more reasonable notion of liberalism, and our (im)possibility theorem provides a somewhat more encouraging conclusion about the compatibility of liberalism with respect for unanimity.

Arrow's (1950; 1951; 1963) impossibility theorem similarly shows that, for $n \geq 2$ and $|\mathcal{X}| \geq 3$, there is no SWF that respects unanimity, independence of irrelevant alternatives (IIA) and non-dictatorship. Although we didn't demand non-dictatorship, it actually follows from respect for justified objections whenever $\gtrsim$ is non-trivial. Like Sen, we dispense with IIA.

A key formal difference between our requirements and those of Sen and Arrow is that theirs restrict how society aggregates preferences *across* preference profiles (minimal liberalism for Sen, IIA and non-dictatorship for Arrow). By contrast, our notion of acceptability only requires satisfactory preference aggregation for each profile, separately.

37. In the surveillance example, Sen's notion of liberalism requires that at least two individuals have some dictatorial power over some aspects of state surveillance!

Appendices

A Standard definitions

This appendix collects definitions of standard order-theoretic concepts used in this paper. Let $\mathcal{A}$ be a non-empty set, and $\geq$ a binary relation on it.

For $a, b \in \mathcal{A}$, we write $a \ngeq b$ iff it is not the case that $b \geq a$, and $a > b$ iff $a \geq b$ and $a \ngeq b$. We say that $a, b \in \mathcal{A}$ are *$\geq$-comparable* iff either $a \geq b$ or $b \geq a$.

$\geq$ is *complete* iff every pair $a, b \in \mathcal{A}$ is $\geq$-comparable, *transitive* iff $a \geq b \geq c$ implies $a \geq c$ for $a, b, c \in \mathcal{A}$, *reflexive* iff $a \geq a$ for any $a \in \mathcal{A}$, and *anti-symmetric* iff $a \geq b \geq a$ implies $a = b$ for $a, b \in \mathcal{A}$. $\geq$ is a *partial order* iff it is transitive, reflexive and anti-symmetric; $(\mathcal{A}, \geq)$ is then called a *poset* (partially ordered set).

A subset $A \subseteq \mathcal{A}$ is called a *$\geq$-chain* iff all of its elements are $\geq$-comparable to each other,[38] and an *$\geq$-anti-chain* iff none of of its elements are $\geq$-comparable to each other.

Given a set $A \subseteq \mathcal{A}$, an element $a \in A$ is *$\geq$-minimal in A* iff there is no $b \in A$ such that $a \geq b \ngeq a$,[39] and a *$\geq$-minimum in A* iff $b \geq a$ for every $b \in A$. Maximal and maximum elements are defined analogously.

An element $b \in \mathcal{A}$ is a *$\geq$-upper bound* of a set $A \subseteq \mathcal{A}$ iff $b \geq a$ for each $a \in A$, and a *$\geq$-minimum $\geq$-upper bound* (or *$\geq$-join,* or *$\geq$-supremum*) iff in addition it is a $\geq$-minimum of the set of $\geq$-upper bounds of A. $\geq$-lower bounds and $\geq$-maximum $\geq$-lower bounds (or $\geq$-meets, or $\geq$-infima) are defined analogously. Note that if $\geq$ is anti-symmetric, then the $\geq$-join and -meet of a set are unique if they exist.

A partially ordered set $(\mathcal{A}, \geq)$ is a *complete lattice* iff every subset of $\mathcal{A}$ has a $\geq$-join and -meet, and simply a *lattice* iff this is true for every two-element subset.

B Suzumura's extension theorem

An extension theorem will play a crucial role in proving Lemma 1 (appendix C), the existence theorem (appendix D), and Lemma 4 (appendix E).

38. Not to be confused with P-chains!

39. When $\geq$ is anti-reflexive, one often sees the equivalent definition 'there is no $b \in A$ such that $a > b$'. The two definitions are not equivalent when $\geq$ fails to be anti-reflexive.

Definition 14. Let $\geq$ and $\geq'$ be binary relations on a set $\mathcal{A}$. $\geq'$ is an *extension* of $\geq$ iff for $a, b \in \mathcal{A}$, $b \geq (>)\ a$ implies $b \geq' (>')\ a$.

Definition 15. A binary relation $\geq$ on a set $\mathcal{A}$ is *Suzumura-consistent* iff for $a_1, \ldots, a_K \in \mathcal{A}$, $a_1 \geq a_2 \geq \cdots \geq a_{K-1} \geq a_K$ implies that either $a_1 \geq a_K$ or a_1, a_K are $\geq$-incomparable.

Transitivity clearly implies Suzumura-consistency, but the converse is false.

Suzumura's extension theorem. A binary relation admits a complete and transitive extension iff it is Suzumura-consistent.

This result is due to Suzumura (1976). A proof may be found in e.g. Bossert and Suzumura (2010, p. 45). We use the theorem directly in the proof of the existence theorem (appendix D). In proving Lemma 1 (appendix C) and Lemma 4 (appendix E), we rely on the following corollary.

Suzumura corollary. Let $\geq$ be a transitive binary relation on a set $\mathcal{A}$, and let $a, b \in \mathcal{A}$ be such that $b \not\geq a$. Then $\geq$ admits a complete and transitive extension $\geq'$ such that $a >' b$.

Proof. Let $\geq^{\triangle}$ be the binary relation on $\mathcal{A}$ such that, for any $c, d \in \mathcal{A}$, $c \geq^{\triangle} d$ iff either (i) $c \geq d$ or (ii) $c = a$ and $d = b$. It suffices to show that $\geq^{\triangle}$ admits a complete and transitive extension to $\mathcal{A}$. So by Suzumura's extension theorem, what we must show is that $\geq^{\triangle}$ is Suzumura-consistent.

To this end, let $a_1 \geq^{\triangle} \cdots \geq^{\triangle} a_K$ in $\mathcal{A}$; we must establish that $a_K \not>^{\triangle} a_1$. Let $\mathcal{I} = \{k \leq K : a_k = a \text{ and } w_{k+1} = b\}$, where $K + 1$ is treated as 1 by convention. If $\mathcal{I}$ is empty, then $a_1 \geq^{\triangle} a_K$ by transitivity of $\geq$. If $\mathcal{I} = \{k\}$, suppose toward a contradiction that $a_K > a_1$; then $b = a_{k+1} \geq a_k = a$ by transitivity of $\geq$, contradicting the hypothesis that $b \not\geq a$. Finally, suppose that $|\mathcal{I}| > 1$. Then there exist $k_1 < k_2$ such that $a_{k_1} = b$, $a_{k_2} = a$, and $a_k \geq a_{k+1}$ for all $k_1 \leq k < k_2$. It follows by transitivity of $\geq$ that $b \geq a$, a contradiction with $b \not\geq a$—thus $|\mathcal{I}| \leq 1$. ■

C Proof of Lemma 1 (§2.1, p. 128)

Fix a $P \subseteq \mathcal{P}$ and a pair $x, y \in \mathcal{X}$ with $x \gtrsim y$. The result is immediate if $x = y$, so suppose that $x > y$.

No P-chain: Suppose that there is no P-chain from x to y. Let $\succeq^{\triangle}$ be the (in general, incomplete) binary relation on $\mathcal{X}$ such that for $z, z' \in \mathcal{X}$, $z \succeq^{\triangle} z'$ iff $z > z'$ and there is a P-chain from z to z'. We seek an upper bound $\succeq''$ of P such that $y \succ'' x$. Equivalently, by the UB characterisation lemma (p. 126), we must show that $\succeq^{\triangle}$ admits a complete and transitive extension $\succeq''$ such that $y \succ'' x$. Since $\succeq^{\triangle}$ is obviously transitive, and $x \not\succeq^{\triangle} y$, this follows from the Suzumura corollary (appendix B).

No strict P-chain: Suppose that there is no strict P-chain from x to y. We seek an upper bound $\succeq''$ of P such that $y \succeq'' x$. If there is no P-chain from x to y, then result follows from the first part of the proof, so assume that there is a P-chain from x to y.

Let $\succeq^{\triangle}$ be the (in general, incomplete) binary relation on $\mathcal{X}$ such that for $z, z' \in \mathcal{X}$, $z \succeq^{\triangle} z'$ iff either (i) $z > z'$ and there is a P-chain from z to z', or (ii) $z' > z$ and (z', z) is part of some P-chain from x to y.[40] The definition of $\succeq^{\triangle}$ may be rephrased as follows: (a) any distinct $z, z' \in \mathcal{X}$ that are not $\gtrsim$-comparable are not $\succeq^{\triangle}$-comparable, (b) for any distinct $z, z' \in \mathcal{X}$ that are $\gtrsim$-comparable, wlog $z > z'$, the following are equivalent: (b1) z, z' are $\succeq^{\triangle}$-comparable, (b2) $z \succeq^{\triangle} z'$, and (b3) there is a P-chain from z to z', and finally (c) for any $z, z' \in \mathcal{X}$ such that $z \succeq^{\triangle} z'$, we have $z' \succeq^{\triangle} z$ if and only if (z, z') is part of some P-chain from x to y.

We claim that any complete and transitive extension $\succeq''$ of $\succeq^{\triangle}$ is an upper bound of P and satisfies $y \succeq'' x$. For the latter, we have $x \succeq'' y$ by (b) since there is a P-chain from x to y; hence (x, y) is part of a P-chain from x to y, whence $y \succeq'' x$ by (c). For the former, we need $\succeq''$ to respect P-chains and strict P-chains in the sense of the UB characterisation lemma (p. 126). Since $\succeq''$ is reflexive (because it lives in $\mathcal{P}$), it respects P-chains of length 1. By (b), $\succeq''$ respects longer P-chains. To show that is also respects *strict* P-chains of length ≥ 2, let $z, z' \in \mathcal{X}$ be such that $z > z'$ and there is a strict P-chain from z to z'. By (b), $z \succeq'' z'$. By (c), we have $z \succ'' z'$ unless (z, z') is part of some P-chain from x to y. But

40. Explicitly, the latter means that there exists a P-chain $(w_l)_{l=1}^{L}$ from x to y such that $w_l = z'$ and $w_{l'} = z$ for some $l < l' \leq L$.

the latter cannot be, for this would produce a strict P-chain from x to y, contradicting our hypothesis. Thus $\succeq''$ is an upper bound of P, and the claim is proved.

It therefore suffices to show that $\succeq^{\Delta}$ admits a complete and transitive extension. By Suzumura's extension theorem (appendix B), this may be achieved by proving that $\succeq^{\Delta}$ is Suzumura-consistent. To that end, fix $w_1, \ldots, w_K \in \mathcal{X}$ such that $w_1 \succeq^{\Delta} \cdots \succeq^{\Delta} w_K$, and suppose toward a contradiction that $w_K \succ^{\Delta} w_1$. Then by definition of $\succeq^{\Delta}$, it must be that $w_K > w_1$, there is a P-chain C_0 from w_K to w_1, and (w_K, w_1) is not part of any P-chain from x to y. We will construct one P-chain C from x to w_K and another P-chain C' from w_1 to y, so that the concatenation $C \cup C_0 \cup C'$ is a P-chain from x to y. Since (w_K, w_1) is part of this P-chain from x to y, this yields the desired contradiction.

To construct C, note that since $w_K > w_1$, there is no P-chain from w_1 to w_K. Let k_1 be the smallest $k \leq K$ such that there is a P-chain from w_l to w_{l+1} for all $k \leq l < K$.[41] Since there is no P-chain from w_1 to w_K, we have $k_1 > 1$. Hence, there is no P-chain from w_{k_1-1} to w_{k_1}. Since $w_{k_1-1} \succeq^{\Delta} w_{k_1}$, it follows from the definition of $\succeq^{\Delta}$ that $w_{k_1} > w_{k_1-1}$ and that (w_{k_1}, w_{k_1-1}) is part of some P-chain from x to y. Let C_1 be the portion of this P-chain from x to w_{k_1}. Moreover, by definition of k_1, there is a (possibly singleton) P-chain C_1' from w_{k_1} to w_K. The concatenation $C = C_1 \cup C_1'$ is a P-chain from x to w_K, as desired.

To construct the second P-chain C', let k_2 be the largest $k \geq 1$ such that there is a P-chain from w_l to w_{l+1} for all $l < k$.[42] Since there is no P-chain from w_1 to w_K, we have $k_2 < k_1$; hence there is no P-chain from w_{k_2} to w_{k_2+1}. Since $w_{k_2} \succeq^{\Delta} w_{k_2+1}$, it follows that $w_{k_2+1} > w_{k_2}$ and (w_{k_2+1}, w_{k_2}) is part of some P-chain from x to y. Let C_2 be the portion of this P-chain from w_{k_2} to y. Moreover, by definition of k_2, there is a (possibly singleton) P-chain C_2' from w_1 to w_{k_2}. The concatenation $C' = C_2 \cup C_2'$ is a P-chain from w_1 to y. ■

41. If this condition fails to hold for every $k < K$, then $k_1 = K$.
42. If this condition fails to hold for every $k > 1$, then $k_2 = 1$.

D Proof that (3) implies (1) in the existence theorem (§2.2.3, p. 134)

For a given $P \subseteq \mathcal{P}$, let $\succeq^\circ$ be the (in general, incomplete) binary relation that satisfies conditions ($\star$)–($\star\star$) in the characterisation theorem (p. 127) for $\gtrsim$-comparable pairs of alternatives, and that does not rank $\gtrsim$-incomparable pairs of alternatives. For each P, $\succeq^\circ$ clearly exists and is unique.

In light of the characterisation theorem (p. 127), property (1) in the existence theorem requires precisely that $\succeq^\circ$ admit a complete and transitive extension (i.e. an extension that lives in $\mathcal{P}$) for any $P \subseteq \mathcal{P}$.[43] Our task is therefore to show that whenever $\gtrsim$ is crown- and diamond-free, $\succeq^\circ$ admits a complete and transitive extension for every $P \subseteq \mathcal{P}$. We will use Suzumura's extension theorem (appendix B).

We first state two lemmata, then use them to show that (3) implies (1). The remainder of this appendix is devoted to proving the lemmata.

Definition 16. A binary relation $\geq$ on a set $\mathcal{A}$ is *weakly transitive* iff for $a, b, c \in \mathcal{A}$, if $a \geq b \geq c$ and a, c are $\geq$-comparable, then $a \geq c$.

$\succeq^\circ$ is weakly transitive provided $\gtrsim$ contains no diamonds:

Lemma 6 (weak transitivity of $\succeq^\circ$)**.** Suppose that $\gtrsim$ is diamond-free. Then $\succeq^\circ$ is weakly transitive for any $P \subseteq \mathcal{P}$.

Weak transitivity is implied by Suzumura-consistency (take $K = 3$).[44] The converse is false in general,[45] but true for $\succeq^\circ$ when $\gtrsim$ has no crowns:

Lemma 7 (Suzumura-consistency of $\succeq^\circ$)**.** Suppose that $\gtrsim$ is crown-free. Then if $\succeq^\circ$ is weakly transitive, it is Suzumura-consistent.

Proof that (3) implies (1). Suppose that $\gtrsim$ is crown- and diamond-free, and fix any $P \subseteq \mathcal{P}$. Since $\gtrsim$ is diamond-free, $\succeq^\circ$ is weakly transitive by Lemma 6. Since $\gtrsim$ is crown-free,

43. The term 'extension' was defined in appendix B.
44. Suzumura-consistency was defined in appendix B.
45. Consider $\mathcal{A} = \{a, b, c, d\}$ and the binary relation $\geq$ such that $a > b > c > d > a$ and no other pairs are $\geq$-comparable. This relation satisfies weak transitivity, but clearly violates Suzumura-consistency.

it follows by Lemma 7 that $\succeq^\circ$ is Suzumura-consistent. Invoking Suzumura's extension theorem (appendix B), we conclude that $\succeq^\circ$ admits a complete and transitive extension. ■

It remains to prove Lemmata 6 and 7. Begin with the former. The role of diamond-freeness is to ensure the following 'crossing' property of decreasing sequences:

Observation 2. Suppose that $\gtrsim$ is diamond-free, and consider $x, y, z \in \mathcal{X}$ with $x > z > y$. Let $(w_k)_{k=1}^K$ be a finite decreasing sequence with $w_1 = x$ and $w_K = y$, and let k' be the last $k < K$ such that $w_{k'} \gtrsim z$. Then $z > w_{k'+1}$.

Proof. Since $w_1 = x > z$, there exist $k < K$ such that $w_k \gtrsim z$, so k' is well-defined. It cannot be that $w_{k'+1} \gtrsim z$ by definition of k'. Nor can it be that $z, w_{k'+1}$ are $\gtrsim$-incomparable, for then $(x, z, w_{k'+1}, y)$ is a diamond: $x \gtrsim z \gtrsim y$ (by hypothesis), $x \gtrsim w_{k'+1} \gtrsim y$ (since $(w_k)_{k=1}^K$ is decreasing from x to y), and $z, w_{k'+1}$ are $\gtrsim$-incomparable. Hence $z > w_{k'+1}$ by anti-symmetry of $\gtrsim$. ■

We will use the following piece of notation: given a set $\mathcal{A}$, a sequence $(a_n)_{n=1}^N$ in $\mathcal{A}$ and an element $b \in \mathcal{A}$, we write $(b) \cup (a_n)_{n=1}^N$ for the sequence $(b, a_1, \ldots, a_N)$, and similarly $(a_n)_{n=1}^N \cup (b)$ for the sequence $(a_1, \ldots, a_N, b)$.

Proof of Lemma 6. Suppose that $\gtrsim$ contains no diamonds, and fix a $P \subseteq \mathcal{P}$ and $x, y, z \in \mathcal{X}$ such that $x \succeq^\circ y \succeq^\circ z$ and x, z are $\succeq^\circ$-comparable. We must show that $x \succeq^\circ z$. This is immediate if x, y, z are not distinct, so suppose that they are. Then by anti-symmetry of $\gtrsim$, there are six cases to check: one for each strict ordering by $\gtrsim$ of x, y, z.

Case 1: $x > y > z$. Since $x \succeq^\circ y$ and $x > y$, by $(\star)$ there is a P-chain from x to y. Similarly there is a P-chain from y to z. The concatenation of these two P-chains is a P-chain from x to z; hence $x \succeq^\circ z$ by $(\star)$.

Case 2: $x > z > y$. Since $x \gtrsim y$ and $x \succeq^\circ y$, there is a P-chain $(w_k)_{k=1}^K$ from x to y by $(\star)$. Let k' be the last $k < K$ for which $w_k \gtrsim z$, so that $w_{k'} \gtrsim z > w_{k'+1}$ by Observation 2. By definition of a P-chain, there is a preference $\succeq$ in P such that $w_{k'} \succeq w_{k'+1}$. It must be that $w_{k'+1} \succeq z$, for otherwise $(z) \cup (w_k)_{k=k'+1}^K$ would be a strict P-chain from z to y, in which case $y \not\succeq^\circ z$ by $(\star\star)$, a contradiction. So we have $w_{k'} \succeq w_{k'+1} \succeq z$, which by

transitivity of $\succeq$ yields $w_{k'} \succeq z$. It follows that $(w_k)_{k=1}^{k'} \cup (z)$ is a P-chain from x to z, so that $x \succeq^\circ z$ by $(\star)$.

Case 3: $y > x > z$. This case is similar to the second. Since $y \gtrsim z$ and $y \succeq^\circ z$, there is a P-chain $(w_k)_{k=1}^{K}$ from y to z by $(\star)$. Let k' be the last $k < K$ for which $w_k \gtrsim x$, so that $w_{k'} \gtrsim x > w_{k'+1}$ by Observation 2. By definition of a P-chain, there is a preference $\succeq$ in P such that $w_{k'} \succeq w_{k'+1}$. It must be that $x \succeq w_{k'}$, for otherwise $(w_k)_{k=1}^{k'} \cup (x)$ would be a strict P-chain from y to x, in which case $x \not\succeq^\circ y$ by $(\star\star)$, a contradiction. So we have $x \succeq w_{k'} \succeq w_{k'+1}$, which by transitivity of $\succeq$ yields $x \succeq w_{k'+1}$. It follows that $(x) \cup (w_k)_{k=k'+1}^{K}$ is a P-chain from x to z, so that $x \succeq^\circ z$ by $(\star)$.

Case 4: $y > z > x$. Suppose toward a contradiction that $x \not\succeq^\circ z$. Then since $z \gtrsim x$, by $(\star\star)$ there exists a strict P-chain from z to x. Since $y \gtrsim z$ and $y \succeq^\circ z$, there is a P-chain from y to z by $(\star)$. Concatenating these two P-chains yields a strict P-chain from y to x, so that $x \not\succeq^\circ y$ by $(\star\star)$, a contradiction.

Case 5: $z > x > y$. This case is similar to the fourth. Suppose toward a contradiction that $x \not\succeq^\circ z$. Then since $z \gtrsim x$, by $(\star\star)$ there exists a strict P-chain from z to x. Since $x \gtrsim y$ and $x \succeq^\circ y$, there is a P-chain from x to y by $(\star)$. Concatenating these two P-chains yields a strict P-chain from z to y, so that $y \not\succeq^\circ z$ by $(\star\star)$, a contradiction.

Case 6: $z > y > x$. Suppose toward a contradiction that $x \not\succeq^\circ z$. Then by $(\star\star)$, there is a strict P-chain $(w_k)_{k=1}^{K}$ from z to x. Let k' be the last $k < K$ for which $w_k \gtrsim y$, so that $w_{k'} \gtrsim y > w_{k'+1}$ by Observation 2. By definition of a P-chain, there is a preference $\succeq$ in P such that $w_{k'} \succeq w_{k'+1}$. Since $(w_k)_{k=1}^{K}$ is a strict P-chain, there are k'' and $\succeq'' \in P$ be such that $w_{k''} \succ'' w_{k''+1}$.

Case 6, sub-case A: $k'' < k'$. It must be that $y \succ w_{k'}$, for otherwise $(w_k)_{k=1}^{k'} \cup (y)$ would be a strict P-chain from y to x, in which case $x \not\succeq^\circ y$ by $(\star\star)$, a contradiction. So we have $y \succ w_{k'} \succeq w_{k'+1}$, which by transitivity of $\succeq$ yields $y \succ w_{k'+1}$. It follows that $(y) \cup (w_k)_{k=k'+1}^{K}$ is a strict P-chain from y to x, so that $x \not\succeq^\circ y$ by $(\star\star)$, a contradiction.

Case 6, sub-case B: $k'' \geq k'$. It must be that $y \succeq w_{k'}$, for otherwise $(w_k)_{k=1}^{k'} \cup (y)$ would be a strict P-chain from y to x, in which case $x \not\succeq^\circ y$ by $(\star\star)$, a contradiction. So we have $y \succeq w_{k'} \succeq w_{k'+1}$, which by transitivity of $\succeq$ yields $y \succeq w_{k'+1}$. It follows that

$(y) \cup (w_k)_{k=k'+1}^{K}$ is a strict P-chain from y to x, so that $x \not\succeq^\circ y$ by $(\star\star)$, a contradiction. ■

It remains to prove Lemma 7.

Definition 17. For a binary relation $\geq$ on a set $\mathcal{A}$, a *weak cycle* is a finite sequence $(a_k)_{k=1}^{K}$ of distinct elements of $\mathcal{A}$ such that a_k, a_{k+1} are $\geq$-comparable for each $1 \leq k \leq K$, where a_{K+1} is understood as a_1.

Note that crowns and diamonds are both examples of weak cycles. Crown-freeness delivers a 'shortcut' property of weak cycles:

Observation 3. Let $\geq$ be a transitive and crown-free binary relation on a set $\mathcal{A}$. Then for any weak cycle $(a_1, \ldots, a_K)$ in $\geq$, (a_1, a_k, a_K) is a weak cycle in $\geq$ for some $1 < k < K$.

Proof. We prove the contra-positive. Suppose that $\geq$ is transitive and that there is a weak cycle $(a_1, \ldots, a_K)$ such that (a_1, a_k, a_K) is not a weak cycle for any $1 < k < K$. In particular, choose $(a_1, \ldots, a_K)$ to be a weak cycle of *minimal length* with this property. Clearly its length K is ≥ 4. We will show that $(a_1, \ldots, a_K)$ is a crown.

There are two cases, $a_1 > a_2$ and $a_1 < a_2$. We consider the former case only; the latter is analogous. It must be that $a_2 < a_3$, for if $a_2 > a_3$ then a_1, a_3 are $\geq$-comparable by transitivity of $\geq$, so $(a_1, a_3, \ldots, a_K)$ is a weak cycle for which (a_1, a_k, a_K) is not a weak cycle for any $3 \leq k < K$, contradicting the minimality of $(a_1, \ldots, a_K)$. Proceeding using the same argument, we conclude that $a_{k-1} > a_k < a_{k+1}$ for every $1 < k < K$ even, where by convention $a_{K+1} \coloneqq a_1$. (This is apparent, separately, for $k < K$ even and for $k = K$.)

Furthermore, it must be that for $1 \leq k' < k'+2 \leq k'' \leq K$, $a_{k'}, a_{k''}$ are $\geq$-incomparable, for if they were comparable then $(a_1, \ldots, a_{k'}, a_{k''}, \ldots, a_K)$ would be a weak cycle for which (a_1, a_k, a_K) is not a weak cycle for any $3 \leq k < K$, contradicting the minimality of $(a_1, \ldots, a_K)$. It follows that $(a_1, \ldots, a_K)$ is a crown. ■

Proof of Lemma 7. Suppose that $\gtrsim$ is crown-free and that $\succeq^\circ$ is weakly transitive. We must show that for every $K \geq 3$, the following claim holds:

$$\text{for any } x_1, \ldots, x_K \in \mathcal{X},\ x_1 \succeq^\circ \cdots \succeq^\circ x_K \succeq^\circ x_1 \text{ implies } x_1 \succeq^\circ x_K. \qquad C(K)$$

We proceed by strong induction on K. In the base case $K = 3$, $C(3)$ is immediate from weak transitivity of $\succeq^\circ$.

Now suppose that $C(K')$ holds for all $K' \leq K - 1$; we will establish $C(K)$. Take any $x_1, \ldots, x_K \in \mathcal{X}$ with $x_1 \succeq^\circ \cdots \succeq^\circ x_K \succeq^\circ x_1$, wlog distinct. Since $\succeq^\circ$ compares all and only $\gtrsim$-comparable pairs of alternatives, $(x_1, \ldots, x_K)$ is a weak cycle in $\gtrsim$. Since $\gtrsim$ (transitive and) crown-free, Observation 3 implies the existence of a $1 < k < K$ such that (x_1, x_k, x_K) is a weak cycle in $\gtrsim$. We consider three cases.

Case 1: $k = 2$. Since $x_2 \succeq^\circ \cdots \succeq^\circ x_K \succeq^\circ x_2$, the inductive hypothesis (in particular, $C(K-1)$) implies that $x_2 \succeq^\circ x_K$. Since $x_1 \succeq^\circ x_2$, and x_1, x_K are $\succeq^\circ$-comparable because they are $\gtrsim$-comparable, it follows by weak transitivity that $x_1 \succeq^\circ x_K$.

Case 2: $k = K - 1$. This case is analogous to the first.

Case 3: $2 < k < K - 1$. Since $x_1 \succeq^\circ \cdots \succeq^\circ x_k$ and x_1, x_k are $\succeq^\circ$-comparable (because they are $\gtrsim$-comparable), the inductive hypothesis (in particular, $C(k)$) implies that $x_1 \succeq^\circ x_k$. Similarly, $x_k \succeq^\circ \cdots \succeq^\circ x_K$, the $\gtrsim$-comparability of x_k, x_K and the inductive hypothesis yield $x_k \succeq^\circ x_K$. Since x_1, x_K are $\succeq^\circ$-comparable (because they are $\gtrsim$-comparable), it follows by weak transitivity that $x_1 \succeq^\circ x_K$. ■

E Proof of Lemma 4 (§2.3, p. 136)

Since $\gtrsim$ is transitive, and x, y are $\gtrsim$-incomparable, the Suzumura corollary (appendix B, p. 158) implies that $\gtrsim$ admits complete and transitive extensions $\succeq', \succeq'' \in \mathcal{P}$ such that $x \succ' y$ and $y \succ'' x$. $\succeq'$ is an upper bound of $\mathcal{P}$ (and hence of any $P \subseteq \mathcal{P}$) because whenever $z \gtrsim (>) w$, we have $z \succeq' (\succ') w$ since $\succeq'$ extends $\gtrsim$; thus the consequent in the definition of '$\succeq' S \succeq$' (p. 119) is satisfied for any $z, w \in \mathcal{X}$ and $\succeq \in \mathcal{P}$. Similarly for $\succeq''$.

F Results for meets

The following observation permits us easily to obtain analogues of the characterisation and existence theorems and the uniqueness proposition for maximum lower bounds.

Definition 18. Let $\geq$ be a binary relation on a set $\mathcal{A}$. The *inverse* of $\geq$, denoted $\leq$, is the binary relation such that for $a, b \in \mathcal{A}$, $a \leq b$ iff $b \geq a$.

Duality observation. Let $S^{\gtrsim}$ denote the single-crossing-dominance relation induced by the primitive order $\gtrsim$ on $\mathcal{X}$, and let $\lesssim$ be the inverse of $\gtrsim$. $S^{\lesssim}$ is the inverse of $S^{\gtrsim}$.

Proof. Simply observe that for any $\succeq, \succeq' \in \mathcal{P}$,

$$\begin{aligned} \succeq' S^{\gtrsim} \succeq \iff & \; x \gtrsim y \text{ and } x \succeq (\succ)\, y \text{ imply } x \succeq' (\succ')\, y \\ \iff & \; y \lesssim x \text{ and } y \succ' (\succeq')\, x \text{ imply } y \succ (\succeq)\, x \iff \succeq S^{\lesssim} \succeq', \end{aligned}$$

which is to say that $S^{\lesssim}$ is the inverse of $S^{\gtrsim}$. ■

It follows that a $S^{\gtrsim}$-maximum $S^{\gtrsim}$-lower bound is precisely a $S^{\lesssim}$-minimum $S^{\lesssim}$-upper bound. In light of this observation, the analogue of the characterisation theorem for meets is immediate:

Definition 19. For a set $P \subseteq \mathcal{P}$ of preferences and two alternatives $y \gtrsim x$ in $\mathcal{X}$, a *reverse P-chain* from y to x is a finite sequence $(w_k)_{k=1}^{K}$ in $\mathcal{X}$ such that

(i) $w_1 = y$ and $w_K = x$,

(ii′) for every $k < K$, $w_k \lesssim w_{k+1}$, and

(iii) for every $k < K$, $w_k \succeq w_{k+1}$ for some $\succeq \in P$.

A *strict reverse P-chain* is a reverse P-chain with $w_k \succ w_{k+1}$ for some $k < K$ and $\succeq \in P$.

Corollary 4 (meet characterisation)**.** For any $P \subseteq \mathcal{P}$, the following are equivalent:

(1′) $\succeq_\star$ is a maximum lower bound of P.

(2′) $\succeq_\star$ is complete and transitive (lives in $\mathcal{P}$), and satisfies: for any $\gtrsim$-comparable $x, y \in \mathcal{X}$, wlog $y \gtrsim x$,

($\star'$) $x \succeq_\star y$ iff there is a reverse P-chain from x to y, and

($\star\star'$) $y \succeq_\star x$ iff there is no strict reverse P-chain from x to y.

Since $\lesssim$ is crown- and diamond-free (complete) iff $\gtrsim$ is, the existence theorem (uniqueness proposition) extends as follows:

Corollary 5 (meet & join existence)**.** The following are equivalent:

(1) Every set of preferences has a minimum upper bound.

(2) Every set of preferences has a maximum lower bound.

(3) $\gtrsim$ is crown- and diamond-free.

Corollary 6 (meet & join uniqueness)**.** The following are equivalent:

(1) Every set of preferences has *at most* one minimum upper bound.

(2) Every set of preferences has *exactly* one minimum upper bound.

(3) Every set of preferences has *at most* one maximum lower bound.

(4) Every set of preferences has *exactly* one maximum lower bound.

(5) $\gtrsim$ is complete.

G Proof of Corollary 1 (§3.2, p. 140)

Equip the space $2^{\mathcal{P}} \setminus \{\varnothing\}$ of non-empty sets of preferences with the S-induced strong set order $\sqsupseteq$. It is easily verified that $\sqsupseteq$ is a partial order. The consensus C is a correspondence $2^{\mathcal{P}} \setminus \{\varnothing\} \Rightarrow \mathcal{X}$, and Proposition 1 (p. 139) says precisely that it is increasing.

Let Π be the set of $P \subseteq \mathcal{P}$ at which $C(P)$ is non-empty. The restriction of C to Π is a non-empty-valued increasing correspondence into $\mathcal{X} \subseteq \mathbf{R}$. It follows by Theorem 2.7 in Kukushkin (2013) that it admits an increasing selection $\phi : \Pi \to \mathcal{X}$.[46]

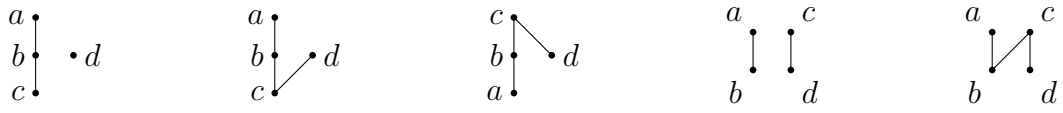
Let

$$\mathcal{U}(P) \coloneqq \{P'' \in \Pi : P'' \sqsupseteq P\} \quad \text{for each non-empty } P \subseteq \mathcal{P},$$

and define the SCF $\psi : 2^{\mathcal{P}} \setminus \{\varnothing\} \to \mathcal{X}$ by

$$\psi(P) \coloneqq \begin{cases} \inf\{\phi(P'') : P'' \in \mathcal{U}(P)\} & \text{if } \mathcal{U}(P) \neq \varnothing \\ \sup \mathcal{X} & \text{otherwise.} \end{cases}$$

46. This step is non-trivial because $\mathcal{X}$ and $\mathcal{P}$ need not be finite, or even countable.



(a) Ball & chain. (b) Hook. (c) Hook. (d) Dumbbells. (e) Saw.

Figure 3: The transitive reductions of the subposets in Definition 20. (Since only the transitive reductions are drawn, the fact that $a > c$ in the ball and chain and the first hook and that $c > a$ in the second hook is implicit.)

The map ψ is well-defined, and really does map into $\mathcal{X}$, because $(\mathcal{X}, \geq)$ is a complete lattice by compactness of $\mathcal{X}$ and the Frink–Birkhoff theorem.[47] Since $\psi = \phi$ on Π (because ϕ is increasing) and ϕ is a selection from C, the SCF ψ respects unanimity.

To see that ψ is monotone, consider $P, P' \in \mathcal{P}$ with $P' \sqsupseteq P$. Observe that $\mathcal{U}(P') \subseteq \mathcal{U}(P)$. If $\mathcal{U}(P')$ is empty then $\psi(P') = \sup \mathcal{X} \geq \psi(P)$ since $\psi(P) \in \mathcal{X}$. If $\mathcal{U}(P)$ is empty then so is $\mathcal{U}(P')$, putting us in the previous case. If neither is empty then

$$\psi(P') = \inf \{\phi(P'') : P'' \in \mathcal{U}(P')\} \geq \inf \{\phi(P'') : P'' \in \mathcal{U}(P)\} = \psi(P). \quad \blacksquare$$

H Elaboration and proof of the (im)possibility theorem (§5.3, p. 154)

In this appendix, we elaborate on the (im)possibility theorem by adding a third equivalence in terms of forbidden subposets (§H.1), then give the proof (§H.2).

H.1 Elaboration of the (im)possibility theorem

It is possible to recast the (im)possibility theorem as a 'forbidden subposet' theorem, much as we did with the existence theorem in appendix I. In particular, $(\mathcal{X}, \gtrsim)$ being a fork or shattered fork is equivalent to the absence from $\gtrsim$ of 4-crowns and diamonds as well as four other four-element subposets: balls-and-chains, hooks, dumbbells and saws. These posets are drawn in Figure 3, and explicitly defined as follows.

Definition 20. Let $\geq$ be a binary relation on a set $\mathcal{A}$, and let (a, b, c, d) be four distinct elements.

47. See e.g. Topkis (1998, Theorem 2.3.1).

(1) (a, b, c, d) is a *ball and chain* iff $a > b > c$ and $a > c$, and no other pair of elements is $\geq$-comparable.

(2) (a, b, c, d) is a *hook* iff either $a > b > c < d$ and $a > c$ or (dually) $a < b < c > d$ and $a < c$, and no other pair of elements is $\geq$-comparable.

(3) (a, b, c, d) is a *pair of dumbbells* iff $a > b$ and $c > d$, and no other pair of elements is $\geq$-comparable.

(4) (a, b, c, d) is a *saw* iff $a > b < c > d$, and no other pair of elements is $\geq$-comparable.

We have the following strengthening of the (im)possibility theorem:

Elaborated (im)possibility theorem. The following are equivalent:

(1) There exists an acceptable SWF.

(2) $(\mathcal{X}, \gtrsim)$ is either a fork or a shattered fork.

(3) $\gtrsim$ is free from 4-crowns, diamonds, balls-and-chains, hooks, dumbbells and saws.

Note that by ruling out saws, we rule out K-crowns for $K \geq 6$ even, since these contains saws.

H.2 Proof of the elaborated (im)possibility theorem

Proof that (1) implies (3). We prove the contra-positive. 4-crowns and diamonds are ruled out by Lemma 5 (p. 153) and the existence theorem (p. 134). For the remainder, it suffices by Lemma 5 to provide, for each forbidden subposet, a pair of preferences $\pi = (\succeq_1, \succeq_2) \in \mathcal{P}^n$ which possess no minimum upper bound $\succeq^\star$ such that $x \succeq^\star (\succ^\star) y$ whenever $x \succeq_i (\succ_i) y$ for $i \in \{1, 2\}$.

Example 1 (pp. 125, 153) does the job for a hook. For a ball and chain (x, y, z, w),

$$z \succ_1 w \succ_1 x \succ_1 y \quad \text{and} \quad y \succ_2 z \succ_2 w \succ_2 x$$

is an example. For a pair of dumbbells (x, y, z, w),

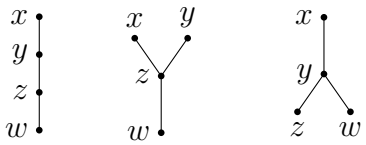
$$x \succ_1 y \succ_1 z \succ_1 w \quad \text{and} \quad z \succ_2 w \succ_2 x \succ_2 y$$

is an example. For a saw (x, y, z, w),

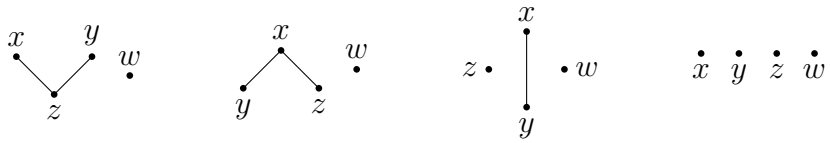
$$w \succ_1 x \succ_1 y \succ_1 z \quad \text{and} \quad y \succ_2 z \succ_2 w \succ_2 x$$

is an example. ■

Proof that (3) implies (2). Suppose that (3) holds; we will argue that $(\mathcal{X}, \gtrsim)$ must be a fork or a shattered fork. By inspection, the only species of four-element subposet permitted by (3) are (three-pronged) fork heads and the following:



(A)



(B)

It is tedious but straightforward to verify that (3) does not allow subposets of *both* types (A) and (B). It is clear that only upward (downward) fork heads are compatible with the second (third) subposet in (A) and with the first (second) subposet in (B).

That leaves two possibilities. The first is that every four-element subposet is an upward (downward) three-pronged fork head, the first subposet in (A), or the second (third) subposet in (A). In that case, $(\mathcal{X}, \gtrsim)$ is evidently an up-fork (down-fork).

The second possibility is that every subposet is an upward (downward) three-pronged fork head, the first (second) subposet in (B), or one of the third and fourth subposets in (B). In that case, $(\mathcal{X}, \gtrsim)$ is a shattered up-fork (down-fork). ■

Proof that (2) implies (1). Suppose that $(\mathcal{X}, \gtrsim)$ is a fork or shattered fork. By Lemma 5 (p. 153), what we must show is that for any $\pi = (\succeq_1, \ldots, \succeq_n) \in \mathcal{P}^n$, there is a minimum upper

bound $\succeq \in \mathcal{P}$ of π that satisfies the following *conditional Pareto property:* $x \succ^\star (\succeq^\star)\ y$ whenever $x \succ_i (\succeq_i)\ y$ for every $i \in \{1, \ldots, n\}$ and there is no (strict) π-chain from y to x. To that end, fix an arbitrary preference profile $\pi \in \mathcal{P}^n$.

Fork: Suppose that $(\mathcal{X}, \gtrsim)$ is a fork, wlog an up-fork. Then $\mathcal{X}$ can be partitioned as $\mathcal{X} = X \cup \{x\} \cup Y$, where $X \cup \{x\}$ is an upward fork head with minimum element x and $\{x\} \cup Y$ is a chain with maximum element x.

Since $\gtrsim$ is crown- and diamond-free, π has minimum upper bounds. Let $\succeq^\star$ be an arbitrary minimum upper bound. Partition X into sets $\{X_k\}_{k=1}^K$ such that $x', x'' \in X$ belong to the same X_k if and only if, for any z in the $\gtrsim$-chain $\{x\} \cup Y$, $x' \succeq^\star (\succ^\star)\ z$ if and only if $x'' \succeq^\star (\succ^\star)\ z$.

Note that, since $X \cup \{x\}$ is an upward fork head with minimum element x, for all $1 \leq k \leq K$, the elements of X_k are mutually $\gtrsim$-incomparable. Moreover, for any $x' \in X$, since $Y \cup \{x, x'\}$ is a $\gtrsim$-chain, all minimum upper bounds of π agree on $Y \cup \{x, x'\}$ by the uniqueness proposition (p. 136). It follows that, for any $(\succeq_k)_{k=1}^K$ where $\succeq_k$ is a preference on X_k for all $1 \leq k \leq K$, there exists a minimum upper bound $\succeq^{\star\star}$ of π that agrees with $\succeq_k$ on X_k for each $k \leq K$.

$\succeq_k$ may clearly be chosen to satisfy the conditional Pareto property on every pair in X_k for each $1 \leq k \leq K$. It thus suffices to show that $\succeq^{\star\star}$ satisfies the conditional Pareto property on pairs x', y' that do not both lie in a single X_k. So suppose that $x' \succ_i (\succeq_i)\ y'$ for every i and that there is no (strict) π-chain from y' to x'; we will deduce that $x' \succ^{\star\star} (\succeq^{\star\star})\ y'$.

There are two cases. First, suppose that one of x' and y' belongs to the $\gtrsim$-chain $\{x\} \cup Y$. Then x' and y' are $\gtrsim$-comparable, so the result follows from properties ($\star$)–($\star\star$) in the characterisation theorem (p. 127). Second, suppose that neither x' nor y' belongs to $\{x\} \cup Y$. Since they do not belong to the same X_k by hypothesis, there is a $z \in \{x\} \cup Y$ such that either (i) $x' \succeq^\star (\succ^\star)\ z$ and $y' \not\succeq^\star (\not\succ^\star)\ z$, or (ii) vice-versa. Since $x' \succ_i (\succeq_i)\ y'$ for every i, it must be (i) that obtains. Therefore $x' \succ^{\star\star} y'$, as desired.

Shattered fork: Suppose that $(\mathcal{X}, \gtrsim)$ is a shattered fork. Then $\mathcal{X}$ may be partitioned as $\mathcal{X} = X \cup Y$, where X is a fork head, wlog an upward fork head, and Y is a set of

isolated elements. Write $x_0 \in X$ for the minimum element of the upward fork head X.

Let $\geq'$ be the binary relation on $\mathcal{X}$ such that, for any $\gtrsim$-comparable $x, y \in \mathcal{X}$, wlog $x \gtrsim y$, $\geq'$ satisfies properties ($\star$)–($\star\star$) in the characterisation theorem (p. 127) and, for any $\gtrsim$-unranked x, y, $x \geq' y$ if and only if $x \succeq_i y$ for every i. Let $\geq''$ be the transitive closure of $\geq'$. Since $\geq''$ is transitive, it admits a complete and transitive extension $\geq^\star$ by Suzumura's extension theorem (appendix B). We will show that $\geq^\star$ is a minimum upper bound of π and that it satisfies the conditional Pareto property.

For the conditional Pareto property, it suffices to show that, for any $x, y \in \mathcal{X}$ such that there is no (strict) π-chain from y to x and further $x \succ_i (\succeq_i) y$ for every i, we have $x >''(\geq'') y$. Note that, if there is no (strict) π-chain from y to x and $x \succ_i (\succeq_i) y$ for every i, then $x \geq' y$. It remains only to show that if there is no π-chain from y to x and $x \succ_i y$ for every i, then $x >'' y$.

So suppose toward a contradiction that there is no π-chain from y to x and that $x \succ_i y$ for every i, and yet $y \geq'' x$. Since $\geq''$ is the transitive closure of $\geq'$, there exists a $\geq'$-decreasing sequence $(w_k)_{k=1}^K$ with $w_1 = y$ and $w_K = x$; we will show that this is impossible. We proceed by strong induction on the length $K \geq 2$ of the sequence. The base case $K = 2$ follows by definition of $\geq'$. For the inductive step, take $K' > 2$, and suppose that there is no $\geq$-decreasing sequence of length K from y to x for any $K < K'$. Toward a contradiction, suppose that there is a $\geq'$-decreasing sequence $(w_k)_{k=1}^{K'}$ of length K' from y to x. There are two cases.

Case 1.1: There is at most one $k' < K'$ such that there is a π-chain from $w_{k'}$ to $w_{k'+1}$. Fix k', and note that $w_k \succeq_i w_{k+1}$ for every i and $k \neq k'$. By hypothesis, we have $x \succ_i y$ for every i. It follows by transitivity that $w_{k'+1} \succ_i w_{k'}$. Since all π-chains have length at most 2, there is no π-chain from $w_{k'}$ to $w_{k'+1}$. But then, since $w_{k'+1} \succ_i w_{k'}$ for every i, it cannot be that $w_{k'} \geq' w_{k'+1}$—a contradiction.

Case 1.2: There are $k_1 < k_2 < K'$ such that there is a π-chain from w_{k_n} to w_{k_n+1} for $n = 1, 2$. Then, since all π-chains end at x_0, $w_{k_1+1} = w_{k_2+1} = x_0$. Then $(w_1, \ldots, w_{k_1}, w_{k_2+1}, \ldots, w_K)$ is a $\geq'$-chain and, since $k_1 < k_2$, it has length $< K'$. This contradicts the inductive hypothesis.

It remains to show that $\geq^\star$ is a minimum upper bound of π. Since $\geq^\star$ is an extension of $\geq''$, it suffices by the characterisation theorem (p. 127) to show that for any $x > y$, $\geq''$ satisfies properties $(\star)$–$(\star\star)$. Since $\geq''$ is the transitive closure of $\geq'$ and $\geq'$ satisfies $(\star)$–$(\star\star)$, we need only show that (a) $y \not\geq'' x$ if there is a strict π-chain from x to y, and that (b) $x \not\geq'' y$ if there is no π-chain from x to y.

Claim (b) is easy: suppose toward a contradiction that there is no π-chain from x to y, and yet $x \geq'' y$. Since $x > y$ by assumption, it follows that $y \succ_i x$ for every i. Then the argument used to show that $\geq^\star$ satisfies the conditional Pareto property applies, yielding a contradiction.

To prove (a), suppose toward a contradiction that there is a strict π-chain from x to y, yet $y \geq'' x$. Then there exists a $\geq'$-decreasing sequence $(w_k)_{k=1}^K$ from $w_1 = y$ to $w_K = x$. We show by strong induction on $K \geq 2$ that this is impossible. The base case $K = 2$ follows by definition of $\geq'$. For the inductive step, take any $K' > 2$, and suppose that for all $K < K'$, there is no $\geq'$-decreasing sequence of length K from y to x. Toward a contradiction, suppose that $(w_k)_{k=1}^{K'}$ is a $\geq'$-decreasing sequence of length K'. There are two cases.

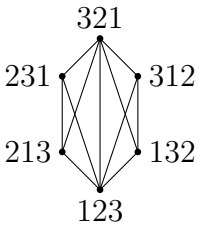
Case 2.1: For some $k < K'$, there is a π-chain from w_k to w_{k+1}. Then, since all π-chains end at x_0, $w_{k+1} = y = x_0$. Then, as above, the induction hypothesis produces a contradiction.

Case 2.2: For every $k < K'$, there is no π-chain from w_k to w_{k+1}. Then since $w_k \geq' w_{k+1}$, we have $w_k \succeq_i w_{k+1}$ for every i. It follows by transitivity that $y \succeq_i x$ for every i. Since all π-chains have length 2, this contradicts the existence of a strict π-chain from x to y. ■

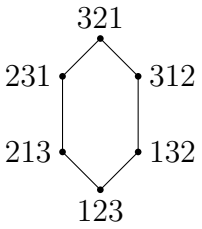
Supplemental appendices

I A characterisation of crown- and diamond-freeness

In this appendix, we provide an alternative way of thinking about crown- and diamond-freeness of $\gtrsim$ by showing that it is equivalent to the absence of *weak cycles* and *chalices* in the *transitive reduction*, or *Hasse diagram*. This provides a novel way to interpret the

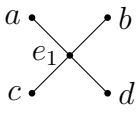


(a) A partial order.

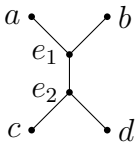


(b) Its transitive reduction.

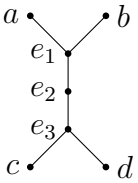
Figure 4: The diagram of a partial order and of its transitive reduction. (In fact, the depicted poset is $(\mathcal{P}^\dagger, S)$, where $\mathcal{P}^\dagger = \{123, 213, 231, 132, 312, 321\}$ is the set of anti-symmetric preferences over alternatives $\mathcal{X} = \{1, 2, 3\}$ ordered by $\geq$ (i.e. $3 > 2 > 1$), and S is single-crossing dominance.)



(a)



(b)



(c)

Figure 5: Chalices.

existence theorem (p. 134). $\mathcal{X}$ is assumed finite for this result.

Definition 21. The *transitive reduction* of a transitive binary relation $\geq$ on a set $\mathcal{A}$ is the smallest binary relation whose transitive closure is $\geq$.[48]

Provided $\mathcal{A}$ is finite, the transitive reduction exists and is unique for any binary relation $\geq$.

A common way to graphically depict a partial order is to draw only its transitive reduction—this is called a *Hasse diagram*. The transitive reduction may be found by deleting, one by one, any edges ('links') that are implied by other edges via transitivity. An example is drawn in Figure 4.

Definition 22. Given an anti-symmetric relation $\geq$ on a set $\mathcal{A}$, a *chalice* is five or more distinct elements $(a, b, e_1, \ldots, e_K, c, d)$ such that $a, b \geq e_1 \geq \cdots \geq e_K \geq c, d$ and a, b and c, d are $\geq$-incomparable. $\geq$ is *chalice-free* iff it contains no chalices.

Some chalices are drawn in Figure 5.

48. The *transitive closure* of a binary relation $\geq$ is the smallest transitive binary relation that contains $\geq$.

Proposition 4 (characterisation of crown- and diamond-freeness)**.** For a partial order $\geq$ on a finite set $\mathcal{A}$, the following are equivalent:

(1) $\geq$ is crown- and diamond-free.

(2) The transitive reduction of $\geq$ is weak-cycle- and chalice-free.[49]

Proposition 4 provides a straightforward graphical interpretation of crown- and diamond-freeness. For example, the partial order drawn in Figure 4a must have a crown or diamond since its transitive reduction in Figure 4b contains a weak cycle. And indeed, it contains several diamonds: $(321, 231, 312, 123)$ is one, and there are three others.

For the proof, we rely on a close cousin of Proposition 4 due to Ball, Pultr and Sichler (2006).

Definition 23. Let $\geq$ be a binary relation on a set $\mathcal{A}$. A 4-crown (a, b, c, d) in $\geq$ *improper* iff there is an $e \in \mathcal{A}$ such that $a, c \geq e \geq b, d$, and *proper* otherwise.

Proposition 5 (Ball, Pultr and Sichler (2006, Proposition 3.1))**.** For a partial order $\geq$ on a finite set $\mathcal{A}$, the following are equivalent:

(1) $\geq$ is free of diamonds, of K-crowns for $K \geq 6$ even, and of proper 4-crowns.

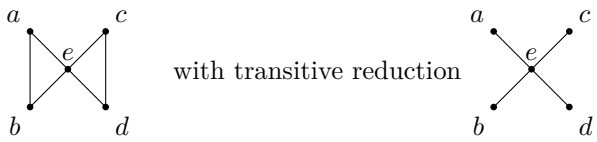
(2) The transitive reduction of $\geq$ is weak-cycle-free.

Proposition 5 implies that improper 4-crowns are consistent with the absence of weak cycles in the transitive reduction.[50] Our Proposition 4 asserts that to avoid improper 4-crowns in $\geq$, it is necessary and sufficient for the transitive reduction to be chalice-free.

Proof of Proposition 4. Suppose that $\geq$ is crown- and diamond-free, and let $\geq^{\top}$ be its transitive reduction. Then $\geq^{\top}$ is weak-cycle-free by Proposition 5, and cannot contain a chalice $(a, c, e_1, \ldots, e_K, b, d)$ since then (a, b, c, d) would be a 4-crown in $\geq$.

49. Weak cycles were defined in appendix D (p. 164).

50. For example, consider the poset



The transitive reduction is weak-cycle free, but (a, b, c, d) is an improper 4-crown.

Suppose that the transitive reduction $\geq^\top$ of $\geq$ is weak-cycle- and chalice-free. By Proposition 5, we need only rule out improper 4-crowns. Suppose toward a contradiction that $\geq$ contains an improper 4-crown (a', b', c', d'). Define a (c) to be the $\geq$-smallest element that is $\geq$-comparable to a' (to c') but $\geq$-incomparable to c' (to a'), and similarly define b and d. Observe that (a, b, c, d) is also an improper 4-crown in $\geq$.

The set

$$E = \{e \in \mathcal{A} : a, c \geq e \geq b, d\}$$

is non-empty since (a, b, c, d) is improper. Furthemore, E is a $\geq$-chain: for if to the contrary $e, e' \in E$ are $\geq$-incomparable, then either (i) $e \vee e'$ and $e \wedge e'$ exist, in which case $(e \vee e', e, e', e \wedge e')$ is a diamond in $\geq$, or (ii) $e \vee e'$ fails to exist, in which case (a, e, c, e') is a proper 4-crown in $\geq$, or (iii) $e \wedge e'$ fails to exist, in which case (e, b, e', d) is a proper 4-crown in $\geq$. We may therefore enumerate E as $e_1 \geq^\top \cdots \geq^\top e_K$. Then $(a, c, e_1, \ldots, e_K, b, d)$ is a chalice in $\geq^\top$, a contradiction. ■

J Non-emptiness is essential in Proposition 2 (§3.3, p. 140)

The following example shows that without the hypothesis that $X(\succeq)$ is non-empty for each $\succeq \in P \cup P'$, the conclusion of Proposition 2 may fail.

Example 8. Consider $\mathcal{X} = \mathbf{N}$, with $\geq$ the usual inequality. Let $\succeq, \succeq' \in \mathcal{P}$ be represented by the utility functions u and u' defined by

$$u(x) := \begin{cases} 1/x & \text{for } x \in \mathcal{X} \text{ odd} \\ 0 & \text{for } x \in \mathcal{X} \text{ even} \end{cases} \quad \text{and} \quad u'(x) := \begin{cases} 0 & \text{for } x \in \mathcal{X} \text{ odd} \\ 1/x & \text{for } x \in \mathcal{X} \text{ even.} \end{cases}$$

Clearly $X(\succeq) = \{1\}$ and $X(\succeq') = \{2\}$. By the uniqueness proposition (p. 136), since $\geq$ is complete, $\{\succeq, \succeq'\}$ has exactly one join and one meet.

We claim that the join and meet are $\gtrsim$ and its inverse $\lesssim$, respectively.[51] To see this, observe that $x + 2 \succ x + 1 \succ' x$ for all $x \in \mathcal{X}$ odd. It follows that for any $x' > x$, $(x', x' - 1, \ldots, x + 1, x)$ is a strict $\{\succeq, \succeq'\}$-chain, and that $(x, x + 1, \ldots, x' - 1, x')$ is a

51. The inverse of a binary relation was defined in appendix F (p. 165).

strict reverse $\{\succeq, \succeq'\}$-chain.[52] It follows by the characterisation theorem (p. 127) that $\gtrsim$ and $\lesssim$ and the join and meet, respectively.

Now let $P \coloneqq \{\succeq', \lesssim\}$ and $P' \coloneqq \{\succeq, \gtrsim\}$. P' dominates P in the S-induced strong set order, so a fortiori in the alternative set order. Since $X(\gtrsim)$ is empty and $X(\lesssim) = \{1\}$, we have $X(P) = \{1, 2\}$ and $X(P') = \{1\}$. $X(P')$ fails to dominate $X(P)$ in the alternative set order: taking $x = 2 \in X(P)$ and $x' = 1 \in X(P')$, there is no y' in $X(P')$ such that $y' \gtrsim x \vee x' = 2$.

References

Anscombe, F. J., and R. J. Aumann. 1963. 'A definition of subjective probability'. *Annals of Mathematical Statistics* 34 (1): 199–205. doi:10.1214/aoms/1177704255.

Arrow, Kenneth J. 1950. 'A difficulty in the concept of social welfare'. *Journal of Political Economy* 58 (4): 328–346. doi:10.1086/256963.

———. 1951. *Social choice and individual values.* New York, NY: Wiley.

———. 1963. *Social choice and individual values.* 2nd. New Haven, CT: Yale University Press.

Ball, Richard N., Aleš Pultr and Jiří Sichler. 2006. 'The mysterious 2-crown'. *Algebra Universalis* 55 (2–3): 213–226. doi:10.1007/s00012-006-1955-3.

Bennett, M. K., and G. Birkhoff. 1994. 'Two families of Newman lattices'. *Algebra Universalis* 32 (1): 115–144. doi:10.1007/BF01190819.

Benoît, Jean-Pierre, and Efe A. Ok. 2007. 'Delay aversion'. *Theoretical Economics* 2 (1): 71–113.

Bossert, Walter, and Kotaro Suzumura. 2010. *Consistency, choice, and rationality.* Cambridge, MA: Harvard University Press.

52. Strict reverse P-chains were defined in appendix F (p. 166).

Casadesus-Masanell, Ramón, Peter Klibanoff and Emre Özdenören. 2000. 'Maxmin expected utility over Savage acts with a set of priors'. *Journal of Economic Theory* 92 (1): 35–65. doi:10.1006/jeth.1999.2630.

Cerreia-Vioglio, Simone, David Dillenberger and Pietro Ortoleva. 2015. 'Cautious expected utility and the certainty effect'. *Econometrica* 83 (2): 639–728. doi:10.3982/ECTA11733.

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci and L. Montrucchio. 2011. 'Uncertainty averse preferences'. *Journal of Economic Theory* 146 (4): 1275–1330. doi:10.1016/j.jet.2011.05.006.

Châtel, Grégory, Vincent Pilaud and Viviane Pons. 2018. 'The weak order on integer posets'. Working paper, 27 Jul 2018. arXiv: 1701.07995v2.

Dekel, Eddie, and Barton L. Lipman. 2008. 'Costly self-control and random self-indulgence'. *Econometrica* 80 (3): 1271–1302. doi:10.3982/ECTA9280.

Duquenne, Vincent, and Ameziane Cherfouh. 1994. 'On permutation lattices'. *Mathematical Social Sciences* 27 (1): 73–89. doi:10.1016/0165-4896(94)00733-0.

Ellsberg, Daniel. 1965. 'Risk, ambiguity, and the Savage axioms'. *Quarterly Journal of Economics* 75 (4): 643–669. doi:10.2307/1884324.

Epstein, Larry G. 1999. 'A definition of uncertainty aversion'. *Review of Economic Studies* 6 (3): 579–608. doi:10.1111/1467-937X.00099.

Ghirardato, Paolo, and Massimo Marinacci. 2002. 'Ambiguity made precise: A comparative foundation'. *Journal of Economic Theory* 102 (2): 251–289. doi:10.1006/jeth.2001.2815.

Gilboa, Itzhak, and David Schmeidler. 1989. 'Maxmin expected utility with non-unique prior'. *Journal of Mathematical Economics* 18 (2): 141–153. doi:10.1016/0304-4068(89)90018-9.

Gray, Robert, and John K. Truss. 2009. 'Cycle-free partial orders and ends of graphs'. *Mathematical Proceedings of the Cambridge Philosophical Society* 146 (3): 535–550. doi:https://doi.org/10.1017/S0305004108002120.

Griggs, Jerrold R., Wei-Tian Li and Linyuan Lu. 2012. 'Diamond-free families'. *Journal of Combinatorial Theory, Series A* 119 (2): 310–322. doi:10.1016/j.jcta.2011.09.002.

Guilbaud, G. T., and P. Rosenstiehl. 1963. 'Analyse algébrique d'un scrutin'. *Mathématiques et Sciences Humaines* 4:9–33.

Gül, Faruk, and Wolfgang Pesendorfer. 2001. 'Temptation and self-control'. *Econometrica* 69 (6): 1403–1435. doi:10.1111/1468-0262.00252.

Horowitz, John K. 1992. 'Comparative impatience'. *Economics Letters* 38 (1): 25–29. doi:10.1016/0165-1765(92)90156-S.

Klibanoff, Peter, Massimo Marinacci and Sujoy Mukerji. 2005. 'A smooth model of decision making under ambiguity'. *Econometrica* 73 (6): 1849–1892. doi:10.1111/j.1468-0262.2005.00640.x.

Kukushkin, Nikolai Seraphimovich. 2013. 'Increasing selections from increasing multifunctions'. *Order* 30 (2): 541–555. doi:10.1007/s11083-012-9260-6.

LiCalzi, Marco, and Arthur F. Veinott. 1992. 'Subextremal functions and lattice programming'. Working paper. doi:10.2139/ssrn.877266.

Loury, Glenn C. 1994. 'Self-censorship in public discourse: A theory of "political correctness" and related phenomena'. *Rationality and Society* 6 (4): 428–461. doi:10.1177/1043463194006004002.

Lu, Linyuan. 2014. 'On crown-free families of subsets'. *Journal of Combinatorial Theory, Series A* 126:216–231. doi:10.1016/j.jcta.2014.04.009.

Maccheroni, Fabio, Massimo Marinacci and Aldo Rustichini. 2006. 'Ambiguity aversion, robustness, and the variational representation of preferences'. *Econometrica* 74 (6): 1447–1498. doi:10.1111/j.1468-0262.2006.00716.x.

Markowsky, George. 1991. 'Permutation lattices revisited'. *Mathematical Social Sciences* 27 (1): 59–72. doi:10.1016/0165-4896(94)00731-4.

Milgrom, Paul, and Chris Shannon. 1994. 'Monotone comparative statics'. *Econometrica* 62 (1): 157–180. doi:10.2307/2951479.

Mirrlees, James A. 1971. 'An exploration in the theory of optimum income taxation'. *Review of Economic Studies* 38 (2): 175–208. doi:10.2307/2296779.

Morris, Stephen. 2001. 'Political correctness'. *Journal of Political Economy* 109 (2): 231–265. doi:10.1086/319554.

Savage, Leonard J. 1954. *The foundations of statistics*. New York, NY: Wiley.

Schmeidler, David. 1989. 'Subjective probability and expected utility without additivity'. *Econometrica* 57 (3): 571–587. doi:10.2307/1911053.

Sen, Amartya. 1970. 'The impossibility of a Paretian liberal'. *Journal of Political Economy* 78 (1): 152–157. doi:10.1086/259614.

Spence, Michael. 1973. 'Job market signaling'. *Quarterly Journal of Economics* 87 (3): 355–374. doi:10.2307/1882010.

Suzumura, Kotaro. 1976. 'Remarks on the theory of collective choice'. *Economica*, New Series, 43 (172): 381–390. doi:10.2307/2553273.

Topkis, Donald M. 1978. 'Minimizing a submodular function on a lattice'. *Operations Research* 26 (2): 305–321. doi:10.1287/opre.26.2.305.

———. 1998. *Supermodularity and complementarity*. Princeton, NJ: Princeton University Press.

Truss, John K. 1996. ‘Betweenness relations and cycle-free partial orders’. *Mathematical Proceedings of the Cambridge Philosophical Society* 119 (4): 631–643. doi:10.1017/S0305004100074478.

von Neumann, John, and Oskar Morgenstern. 1947. *Theory of games and economic behavior.* 2nd. Princeton, NJ: Princeton University Press.

Yaari, Menahem E. 1969. ‘Some remarks on measures of risk aversion and on their uses’. *Journal of Economic Theory* 1 (3): 315–329. doi:10.1016/0022-0531(69)90036-2.

Yanagimoto, Takemi, and Masashi Okamoto. 1969. ‘Partial orderings of permutations and monotonicity of a rank correlation statistic’. *Annals of the Institute of Statistical Mathematics* 21 (1): 489–506. doi:10.1007/BF02532273.