The converse envelope theorem*

Ludvig Sinander†

Abstract

I prove an envelope theorem with a converse: the envelope formula is *equivalent* to a first-order condition. Like Milgrom and Segal's (2002) envelope theorem, my result requires no structure on the choice set. I use the converse envelope theorem to extend to general outcomes and preferences the canonical result in mechanism design that any increasing allocation is implementable, and apply this to selling information.

Keywords: Envelope theorem, first-order condition, mechanism design

1 Introduction

Envelope theorems are a key tool of economic theory, with important roles in consumer theory, mechanism design and dynamic optimisation. In blueprint form, an envelope theorem gives conditions under which optimal decision-making implies that the *envelope formula* holds.

In textbook accounts,[1] the envelope theorem is typically presented as a consequence of the first-order condition. The modern envelope theorem of Milgrom and Segal (2002), however, applies in an abstract setting in which the first-order condition is typically not even well-defined. These authors therefore rejected the traditional intuition and developed a new one.

In this paper, I re-establish the intuitive link between the envelope formula and the first-order condition. I introduce an appropriate generalised first-order condition that is well-defined in the abstract environment of Milgrom and Segal (2002), then prove an envelope theorem with a converse: my generalised first-order condition is

*I am grateful to Eddie Dekel, Alessandro Pavan and Bruno Strulovici for their guidance and support. This work has profited from the close reading and insightful comments of Gregorio Curello, Eddie Dekel, Roberto Saitto, Quitzé Valenzuela-Stookey, Alessandro Lizzeri (the editor) and four anonymous referees, and from comments by Piotr Dworczak, Matteo Escudé, Daniel Gottlieb, Elliot Lipnowski, Benny Moldovanu, Ilya Segal and audiences at Caltech, Northwestern, Oxford, the Bonn Winter Theory Workshop, the Kansas Workshop in Economic Theory and the Southeast Theory Festival.

†Department of Economics, University of Oxford and Nuffield College.

[1] E.g. Mas-Colell, Whinston and Green (1995, §M.L).

equivalent to the envelope formula. This validates the habitual interpretation of the envelope formula as 'local optimality', and clarifies our understanding of the envelope theorem.

The converse envelope theorem proves useful for mechanism design. I use it to establish that the implementability of all increasing allocations, a canonical result when outcomes are drawn from an interval of $\mathbf{R}$, remains valid when outcomes are abstract. I apply this result to the problem of selling information (distributions of posteriors).

The setting is simple: an agent chooses an action x from a set $\mathcal{X}$ to maximise $f(x,t)$, where $t \in [0,1]$ is a parameter. The set $\mathcal{X}$ need not have any structure. A *decision rule* is a map $X : [0,1] \to \mathcal{X}$ that assigns an action $X(t)$ to each parameter t. A decision rule X is associated with a *value function* $V_X(t) \coloneqq f(X(t),t)$, and is called *optimal* iff $V_X(t) = \max_{x \in \mathcal{X}} f(x,t)$ for every parameter t.

The modern envelope theorem of Milgrom and Segal (2002) states that, under a regularity assumption on f, any optimal decision rule X induces an absolutely continuous value function V_X which satisfies the *envelope formula*

$$V_X'(t) = f_2(X(t),t) \quad \text{for a.e. } t \in (0,1).$$

The familiar intuition is as follows. The derivative of the value V_X is

$$V_X'(t) = \frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\bigg|_{m=0} + f_2(X(t),t),$$

where the first term is the indirect effect via the induced change of the action, and the second term is the direct effect. Since X is optimal, it satisfies the first-order condition $\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\big|_{m=0} = 0$, which yields the envelope formula. Indeed, a decision rule X satisfies the envelope formula *if and only if* it satisfies the first-order condition for a.e. $t \in (0,1)$.

The trouble with this intuition is that since the action set $\mathcal{X}$ is abstract (with no linear or topological structure), the derivative $\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\big|_{m=0}$ is ill-defined in general.

To restore the equivalence of the envelope formula and first-order condition, I first introduce a generalised first-order condition that is well-defined in the abstract environment. The *outer first-order condition* is the following 'integrated' variant of the classical first-order condition:

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m),s)\mathrm{d}s\bigg|_{m=0} = 0 \quad \text{for all } r,t \in (0,1).$$

I then prove an envelope theorem with a converse: under a regularity assumption on f, a decision rule X satisfies the envelope formula *if and only if* it satisfies the outer first-order condition and induces an absolutely continuous value function V_X. The 'only if' part is a novel *converse* envelope theorem.

In §4, I apply the converse envelope theorem to mechanism design. There is an agent with preferences over outcomes $y \in \mathcal{Y}$ and payments $p \in \mathbf{R}$. Her preferences are indexed in 'single-crossing' fashion by $t \in [0, 1]$, and this taste parameter is privately known to her. A canonical result is that if $\mathcal{Y}$ is an interval of $\mathbf{R}$, then all (and only) increasing allocations $Y : [0, 1] \to \mathcal{Y}$ can be implemented incentive-compatibly by some payment schedule $P : [0, 1] \to \mathbf{R}$.

I use the converse envelope theorem to extend this result to a large class of ordered outcome spaces $\mathcal{Y}$, maintaining general (non-quasi-linear) preferences. The argument runs as follows: fix an increasing allocation $Y : [0, 1] \to \mathcal{Y}$. To implement it, choose a payment schedule $P : [0, 1] \to \mathbf{R}$ to make the envelope formula hold. Then by the converse envelope theorem, the outer first-order condition is satisfied, which means intuitively that (Y, P) is *locally* incentive-compatible. The single-crossing property of preferences ensures that this translates into global incentive-compatibility.

I apply this implementability theorem to study the sale of information. The result implies that any Blackwell-increasing information allocation is implementable. I argue further that if consumers can share their information with each other, then *only* Blackwell-increasing allocations are implementable.

1.1 Related literature

Envelope theorems entered economics via the theories of the consumer and of the firm (Hotelling, 1932; Roy, 1947; Shephard, 1953), were systematised by Samuelson (1947) under 'classical' assumptions, and were developed in greater generality by e.g. Danskin (1966, 1967), Silberberg (1974) and Benveniste and Scheinkman (1979). Milgrom and Segal (2002) pointed out that classical-type assumptions were extraneous, and proved an envelope theorem without them. Subsequent refinements were obtained by e.g. Morand, Reffett and Tarafdar (2015) and Clausen and Strub (2020).[2] 'Converse' envelope theorems are almost absent from this literature, but appear in textbook presentations (e.g. Mas-Colell, Whinston & Green, 1995, §M.L).

The outer first-order condition appears to be novel. It bears no clear relationship to any of the standard derivatives for non-smooth functions.

[2] See also Oyama and Takenawa (2018).

2 Setting and background

In this section, I introduce the environment, the Milgrom–Segal (2002) envelope theorem, and the classical envelope theorem and converse.

Notation. We will be working with the unit interval $[0,1]$, equipped with the Lebesgue σ-algebra and the Lebesgue measure. The Lebesgue integral will be used throughout. For $r < t$ in $[0,1]$, we will write $\int_r^t$ for the integral over $[r,t]$, and $\int_t^r$ for $-\int_r^t$. $\mathcal{L}^1$ will denote the space of integrable functions $[0,1] \to \mathbf{R}$, i.e. those that are measurable and have finite integral. We will write f_i for the derivative of a function f with respect to its ith argument. Some important definitions and theorems are collected in appendix A, including Lebesgue's fundamental theorem of calculus and the Vitali convergence theorem.

2.1 Setting

An agent chooses an action x from an arbitrary set $\mathcal{X}$. Her objective is $f(x,t)$, where $t \in [0,1]$ is a parameter (or 'type').[3]

Definition 1. A family $\{\phi_x\}_{x \in \mathcal{X}}$ of functions $[0,1] \to \mathbf{R}$ is *absolutely equi-continuous* iff the family of functions

$$\left\{ t \mapsto \sup_{x \in \mathcal{X}} \left| \frac{\phi_x(t+m) - \phi_x(t)}{m} \right| \right\}_{m>0}$$

is uniformly integrable.[4]

Our only assumptions will be that the objective varies smoothly, and (uniformly) not too erratically, with the parameter.

Basic assumptions. $f(x,\cdot)$ is differentiable for every $x \in \mathcal{X}$, and the family $\{f(x,\cdot)\}_{x \in \mathcal{X}}$ is absolutely equi-continuous.

Remark 1. An easy-to-check sufficient condition for absolute equi-continuity is as follows: $f(x,\cdot)$ is absolutely continuous for each $x \in \mathcal{X}$, and there is an $\ell \in \mathcal{L}^1$ such that $|f_2(x,t)| \leq \ell(t)$ for all $x \in \mathcal{X}$ and $t \in (0,1)$. (This is the assumption that Milgrom and Segal (2002) use in their envelope theorem.) An even stronger sufficient condition is that f_2 be bounded.

[3] If instead the parameter lives in a normed vector space, then the analysis applies unchanged to path derivatives (as Milgrom and Segal (2002, footnote 7) point out).

[4] The name 'absolute equi-continuity' is inspired by the AC–UI lemma in appendix A, which states that absolute continuity of a continuous ϕ is equivalent to uniform integrability of the 'divided-difference' family $\{t \mapsto [\phi(t+m) - \phi(t)]/m\}_{m>0}$. As the term suggests, an absolutely equi-continuous family is equi-continuous, and its members are absolutely continuous functions; this is proved in appendix B.

Example 1. Let $\mathcal{X} = [0,1]$ and $f(x,t) = xt$. The basic assumptions are satisfied since $f_2(x,t) = x$ exists and is bounded. $\diamond$

A *decision rule* is a map $X : [0,1] \to \mathcal{X}$ that prescribes an action for each type. The payoff of type t from following decision rule X is denoted $V_X(t) \coloneqq f(X(t),t)$.

Definition 2. A decision rule X satisfies the *envelope formula* iff

$$V_X(t) = V_X(0) + \int_0^t f_2(X(s),s)\mathrm{d}s \quad \text{for every } t \in [0,1].$$

Equivalently (by Lebesgue's fundamental theorem of calculus), X satisfies the envelope formula iff V_X is absolutely continuous and

$$V_X'(t) = f_2(X(t),t) \quad \text{for a.e. } t \in (0,1).$$

A decision rule X is called *optimal* iff at every parameter $t \in [0,1]$, $X(t)$ maximises $f(\cdot,t)$ on $\mathcal{X}$. The modern envelope theorem is as follows:

Milgrom–Segal envelope theorem. Under the basic assumptions, if X is optimal, then it satisfies the envelope formula.

This follows from the main theorem (§3.2 below), so no proof is necessary. It is actually a slight refinement of Theorem 2 in Milgrom and Segal (2002), as these authors impose the sufficient condition in Remark 1 rather than absolute equi-continuity.

Example 1 (continued)**.** The envelope formula requires that $X(t)t = \int_0^t X$ for every $t \in [0,1]$, or equivalently $X(t) = t^{-1}\int_0^t X$ for all $t \in (0,1]$. Thus the decision rules that satisfy the envelope formula are precisely those that are constant on $(0,1]$. This includes all optimal decision rules (which set $X = 1$ on $(0,1]$), as well as anti-optimal ones (which choose 0 on $(0,1]$). $\diamond$

2.2 Classical envelope theorem and converse

The textbook version of the envelope theorem, which has a natural and intuitive converse, holds under additional topological and convexity assumptions.

Classical assumptions. The action set $\mathcal{X}$ is a convex subset of $\mathbf{R}^n$, the action derivative f_1 exists and is bounded, and only Lipschitz continuous decision rules X are considered.

The classical assumptions are strong. Most glaringly, the Lipschitz condition rules out important decision rules in many applications. In the canonical auction setting, for instance, the revenue-maximising mechanism is discontinuous (Myerson, 1981).[5]

Example 1 (continued)**.** $\mathcal{X} = [0,1]$ is a convex subset of $\mathbf{R}$, and $f_1(x,t) = t$ exists and is bounded. If we restrict attention to Lipschitz continuous decision rules $X : [0,1] \to [0,1]$, then the classical assumptions are satisfied. ◇

Given a Lipschitz continuous decision rule X, suppose that type t considers taking the action $X(t+m)$ intended for another type. The map $m \mapsto f(X(t+m),t)$ is differentiable a.e. under the classical assumptions,[6] so we may define a first-order condition:

Definition 3. A decision rule X satisfies the *first-order condition a.e.* iff

$$\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\bigg|_{m=0} = 0 \quad \text{for a.e. } t \in (0,1).$$

The first-order condition a.e. requires that almost no type t can secure a first-order payoff increase (or decrease) by choosing an action $X(t+m)$ intended for a nearby type $t+m$. It does *not* say that there are no nearby *actions* that do better (or worse).

Classical envelope theorem and converse. Under the basic and classical assumptions, a Lipschitz continuous decision rule satisfies the first-order condition a.e. iff it satisfies the envelope formula.

The proof, given in appendix G, shows that the envelope formula demands precisely that $V_X'(t) = f_2(X(t),t)$ for a.e. $t \in (0,1)$, which is equivalent to the first-order condition a.e. by inspection of the differentiation identity

$$V_X'(t) = \frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\bigg|_{m=0} + f_2(X(t),t).$$

Example 1 (continued)**.** A Lipschitz continuous decision rule X is differentiable a.e., so satisfies the first-order condition a.e. iff

$$\frac{\mathrm{d}}{\mathrm{d}m} X(t+m)t\bigg|_{m=0} = X'(t)t = 0 \quad \text{for a.e. } t \in (0,1).$$

This requires that $X' = 0$ a.e. We saw that the envelope formula demands that X be constant on $(0,1]$. For Lipschitz continuous decision rules X, both conditions are equivalent to constancy on all of $[0,1]$. ◇

[5]Even when the classical assumptions are relaxed as much as possible, unless f is trivial, X still has to satisfy a strong continuity requirement. See appendix G.

[6]Since $f(\cdot,t)$ is differentiable, and X is differentiable a.e. since it is Lipschitz continuous.

3 Main theorem

In this section, I define the outer first-order condition and state my envelope theorem and converse.

3.1 The outer first-order condition

Without the classical assumptions (§2.2), the 'imitation derivative'

$$\left.\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m), t)\right|_{m=0}$$

need not exist, in which case the first-order condition is ill-defined. To circumvent this problem, we require a novel first-order condition.

Definition 4. A decision rule X satisfies the *outer first-order condition* iff

$$\left.\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s\right|_{m=0} = 0 \quad \text{for all } r, t \in (0,1).$$

As an intuitive motivation, suppose that types $s \in [r,t]$ deviate by choosing $X(s+m)$ rather than $X(s)$. The aggregate payoff to such a deviation is $\int_r^t f(X(s+m), s)\mathrm{d}s$, and the outer first-order condition says (loosely) that local deviations of this kind are collectively unprofitable.

Example 1 (continued)**.** For any decision rule X that is a.e. constant at some $k \in [0,1]$, the outer first-order condition holds:

$$\left.\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t X(s+m)s\mathrm{d}s\right|_{m=0} = \left.\frac{\mathrm{d}}{\mathrm{d}m} k \int_r^t s\mathrm{d}s\right|_{m=0} = 0 \quad \text{for all } r, t \in (0,1).$$

Conversely, any decision rule that is not constant a.e. violates the outer first-order condition. ◇

As we shall see, the outer first-order condition is well-defined even when the classical assumptions fail. When they do hold, the outer first-order condition coincides with the first-order condition a.e.:

Housekeeping lemma. Under the basic and classical assumptions, the outer first-order condition is equivalent to the first-order condition a.e.

Proof. Fix a Lipschitz continuous decision rule $X : [0,1] \to \mathcal{X}$. The family

$$\left\{ t \mapsto \frac{f(X(t+m), t) - f(X(t), t)}{m} \right\}_{m>0}$$

is convergent a.e. as $m \downarrow 0$ by the classical assumptions, and is uniformly integrable by Lemma 4 in appendix F. Hence by the Vitali convergence theorem, for any $r, t \in (0, 1)$,

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s \bigg|_{m=0} = \int_r^t \frac{\mathrm{d}}{\mathrm{d}m} f(X(s+m), s)\bigg|_{m=0} \mathrm{d}s.$$

The left-hand side (right-hand side) is zero for all $r, t \in (0, 1)$ iff the outer first-order condition (first-order condition a.e.) holds.[7] ■

The term 'outer' is inspired by this argument. By taking the differentiation operator outside the integral, we change nothing in the classical case, and ensure existence beyond the classical case.

As its name suggests, the outer first-order condition is necessary (but not sufficient) for optimality. The following is proved in appendix E:

Necessity lemma. Under the basic assumptions, any optimal decision rule X satisfies the outer first-order condition, and has $V_X(t) := f(X(t), t)$ absolutely continuous.

3.2 Envelope theorem and converse

My main result characterises the envelope formula in terms of the outer first-order condition.

Envelope theorem and converse. Under the basic assumptions, for a decision rule $X : [0, 1] \to \mathcal{X}$, the following are equivalent:

(1) X satisfies the outer first-order condition

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s \bigg|_{m=0} = 0 \quad \text{for all } r, t \in (0, 1),$$

and $V_X(t) := f(X(t), t)$ is absolutely continuous.

(2) X satisfies the envelope formula

$$V_X(t) = V_X(0) + \int_0^t f_2(X(s), s)\mathrm{d}s \quad \text{for every } t \in [0, 1].$$

The implication (1) $\implies$ (2) is an envelope theorem with weak (purely local) assumptions; the Milgrom–Segal and classical envelope theorems in §2 are corollaries.

[7] For the right-hand side, this relies on the following basic fact (e.g. Proposition 2.23(b) in Folland (1999)): for $\phi \in \mathcal{L}^1$, we have $\phi = 0$ a.e. iff $\int_r^t \phi = 0$ for all $r, t \in (0, 1)$.

The implication (2) $\implies$ (1) is the converse envelope theorem, which entails the classical converse envelope theorem in §2.2.

The absolute-continuity-of-V_X condition in (1) ensures that $f(X(\cdot), t)$ does not behave too erratically near t. A characterisation of this property is provided in appendix D.

Example 1 (continued)**.** We saw that a decision rule satisfies the envelope formula iff it is constant on $(0,1]$ (p. 5), and satisfies the outer first-order condition iff it is constant a.e. (p. 7). Thus the envelope formula implies the outer first-order condition. For the other direction, observe that an a.e. constant X for which $V_X(t) = X(t)t$ is (absolutely) continuous must in fact be constant on $(0,1]$, though not necessarily at zero. $\diamondsuit$

In the classical case (§2.2), our proof relied on the differentiation identity

$$V_X'(t) = \frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m), t)\bigg|_{m=0} + f_2(X(t), t),$$

or (rearranged and integrated)

$$\int_r^t \frac{\mathrm{d}}{\mathrm{d}m} f(X(s+m), s)\bigg|_{m=0} \mathrm{d}s = V_X(t) - V_X(r) - \int_r^t f_2(X(s), s)\mathrm{d}s.$$

To pursue an analogous proof, we require an 'outer' version of this identity in which differentiation and integration are interchanged on the left-hand side. The following lemma, proved in appendix C, does the job.

Identity lemma. Under the basic assumptions, if V_X is absolutely continuous, then for all $r, t \in (0,1)$,

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s\bigg|_{m=0} = V_X(t) - V_X(r) - \int_r^t f_2(X(s), s)\mathrm{d}s, \qquad (\mathcal{I})$$

where both sides are well-defined.

The left-hand side of $(\mathcal{I})$ is zero for all $r, t \in (0,1)$ iff the outer first-order condition holds. The right-hand side is zero for all $r, t \in (0,1)$ iff the envelope formula holds.[8] Therefore:

Proof of the envelope theorem and converse. Suppose that the outer first-order condition holds and that V_X is absolutely continuous. Then the identity lemma applies, so the outer first-order condition implies the envelope formula.

[8]For the 'only if' part, if right-hand side is zero for all $r, t \in (0,1)$, then it is zero for all $r, t \in [0,1]$ since V_X and the integral are continuous, yielding the envelope formula.

Suppose that the envelope formula holds. Then V_X is absolutely continuous by Lebesgue's fundamental theorem of calculus. Hence the identity lemma applies, so the envelope formula implies the outer first-order condition. ■

4 Application to mechanism design

A key result in mechanism design is that, provided the agent's preferences are 'single-crossing', all and only increasing allocations are implementable. While the 'only' part is straightforward, the 'all' part has substance. Existing theorems of this sort require that outcomes be drawn from an interval of $\mathbf{R}$ or that the agent have quasi-linear preferences.

In this section, I use the converse envelope theorem to extend this result to abstract spaces of outcomes, without requiring quasi-linearity. I then apply it to the problem of selling information, showing that all (and only) Blackwell-increasing information allocations are implementable (and robust to collusion).

4.1 Environment and existing results

There is a partially ordered set $\mathcal{Y}$ of outcomes. A single agent has preferences over outcomes $y \in \mathcal{Y}$ and payments $p \in \mathbf{R}$ represented by $f(y, p, t)$, where the type $t \in [0, 1]$ is privately known to the agent.[9] We assume that $f(y, \cdot, t)$ is strictly decreasing and onto $\mathbf{R}$ for all $y \in \mathcal{Y}$ and $t \in [0, 1]$.

A *direct mechanism* is a pair of maps $Y : [0, 1] \to \mathcal{Y}$ and $P : [0, 1] \to \mathbf{R}$ that assign an outcome and a payment to each type. A direct mechanism (Y, P) is called *incentive-compatible* iff no type strictly prefers the outcome–payment pair designated for another type:

$$f(Y(t), P(t), t) \geq f(Y(r), P(r), t) \quad \text{for all } r, t \in [0, 1].$$

By a revelation principle, it is without loss of generality to restrict attention to incentive-compatible direct mechanisms. An allocation $Y : [0, 1] \to \mathcal{Y}$ is called *implementable* iff there is a payment schedule $P : [0, 1] \to \mathbf{R}$ such that (Y, P) is incentive-compatible.[10] An *increasing* allocation is one that provides higher types with larger outcomes (in the partial order on $\mathcal{Y}$).

Preferences f are called *single-crossing* iff higher types are more willing to pay to increase $y \in \mathcal{Y}$. The details of how this is formalised vary from paper to paper. We are interested in the following type of result:

[9] All of the analysis carries over to the case of multiple agents with independent types.
[10] Adding an individual rationality constraint does not change our results below.

Theorem schema. If $\mathcal{Y}$ and f are 'regular' and f is 'single-crossing', then any increasing allocation is implementable.

The first result of this kind was obtained by Mirrlees (1976) and Spence (1974) under the assumptions that $\mathcal{Y}$ is an interval of $\mathbf{R}$ and that f has the quasi-linear form $f(y, p, t) = h(y, t) - p$. Maintaining quasi-linearity, the result was extended to multi-dimensional Euclidean $\mathcal{Y}$ by Matthews and Moore (1987) and García (2005),[11] and may be further extended to arbitrary $\mathcal{Y}$ via a standard argument. (That argument relies critically on quasi-linearity; see supplemental appendix K.) With $\mathcal{Y}$ an interval of $\mathbf{R}$, the result was obtained without quasi-linearity by Guesnerie and Laffont (1984) under classical assumptions,[12] and by Nöldeke and Samuelson (2018) assuming only that f is (jointly) continuous.

I shall extend the result to a wide class of outcome spaces $\mathcal{Y}$, without imposing quasi-linearity. I formulate notions of 'regularity' and 'single-crossing' in the next section, then establish the implementability of increasing allocations in §4.3.

4.2 Regularity and single-crossing

Recall that a subset $\mathcal{C} \subseteq \mathcal{Y}$ is called a *chain* iff it is totally ordered.

Definition 5. The outcome space $\mathcal{Y}$ is *regular* iff it is order-dense-in-itself, countably chain-complete and chain-separable.[13]

In words, $\mathcal{Y}$ must be 'rich' (first two assumptions) and 'not too large' (final assumption). Many important spaces enjoy these properties, including $\mathbf{R}^n$ with the usual (product) order, the space of finite-expectation random variables (on some probability space) ordered by 'a.s. smaller', and the space of distributions of posteriors updated from a given prior ordered by Blackwell informativeness. I prove these assertions and give further examples in supplemental appendix L.

Definition 6. The payoff f is *regular* iff (a) the type derivative f_3 exists and is bounded, and $f_3(y, \cdot, t)$ is continuous for each $y \in \mathcal{Y}$ and $t \in [0, 1]$, and (b) for every chain $\mathcal{C} \subseteq \mathcal{Y}$, f is jointly continuous on $\mathcal{C} \times \mathbf{R} \times [0, 1]$ when $\mathcal{C}$ has the relative topology inherited from the order topology on $\mathcal{Y}$.[14,15]

[11] Results of this type have been used to study sequential screening (e.g. Courty and Li (2000), Battaglini (2005), Eső and Szentes (2007), and Pavan, Segal and Toikka (2014)).

[12] These authors restricted attention to piecewise continuously differentiable allocations; Milgrom (2004, Theorem 4.2) generalised to piecewise absolutely continuous allocations.

[13] A set $\mathcal{A}$ partially ordered by $\lesssim$ is *order-dense-in-itself* iff for any $a < a'$ in $\mathcal{A}$, there is a $b \in \mathcal{A}$ such that $a < b < a'$. $B \subseteq \mathcal{A}$ is *order-dense* in $C \subseteq \mathcal{A}$ iff for any $c < c'$ in C, there is a $b \in B$ such that $c \lesssim b \lesssim c'$. $\mathcal{A}$ is *chain-separable* iff for each chain $C \subseteq \mathcal{A}$, there is a countable set $B \subseteq \mathcal{A}$ that is order-dense in C. $\mathcal{A}$ is *countably chain-complete* iff every countable chain in $\mathcal{A}$ with a lower (upper) bound in $\mathcal{A}$ has an infimum (a supremum) in $\mathcal{A}$.

[14] The *order topology* on $\mathcal{Y}$ is the one generated by the open order rays $\{y' \in \mathcal{Y} : y' < y\}$ and $\{y' \in \mathcal{Y} : y < y'\}$ for each $y \in \mathcal{Y}$, where $<$ denotes the strict part of the order on $\mathcal{Y}$.

[15] It is sufficient, but unnecessarily strong, to assume joint continuity on $\mathcal{Y} \times \mathbf{R} \times [0, 1]$.

The joint continuity requirement corresponds to Nöldeke and Samuelson's (2018) regularity assumption. By demanding in addition that the type derivative exist and be bounded, I ensure that when this model is embedded in the general setting of §2.1 by letting $\mathcal{X} \coloneqq \mathcal{Y} \times \mathbf{R}$, the basic assumptions are satisfied. The converse envelope theorem is thus applicable.[16]

It remains to formalise 'single-crossing', the idea that higher types are more willing to pay to increase $y \in \mathcal{Y}$. Under the classical assumptions, this is captured by the *Spence–Mirrlees condition,* which demands that for any increasing $Y : [0,1] \to \mathcal{Y}$ and any $P : [0,1] \to \mathbf{R}$ (both Lipschitz continuous), for any type $s \in (0,1)$, the marginal gain to mimicking

$$\frac{\mathrm{d}}{\mathrm{d}m} f(Y(s+m), P(s+m), s+n)\bigg|_{m=0}$$

be single-crossing in n.[17,18] To extend this definition beyond the classical case to general outcomes $\mathcal{Y}$ (and non-Lipschitz mechanisms (Y, P)), I replace the (typically ill-defined) marginal mimicking gain with its 'outer' version:

Definition 7. f satisfies the *(strict) outer Spence–Mirrlees condition* iff for any increasing $Y : [0,1] \to \mathcal{Y}$, any $P : [0,1] \to \mathbf{R}$ and any $r < t$ in $(0,1)$,

$$n \mapsto \frac{\overline{\mathrm{d}}}{\overline{\mathrm{d}}m} \int_r^t f(Y(s+m), P(s+m), s+n)\mathrm{d}s\bigg|_{m=0}$$

is (strictly) single-crossing, where $\overline{\mathrm{d}}/\overline{\mathrm{d}}m$ denotes the upper derivative.[19]

The difference from the classical Spence–Mirrlees condition is merely technical: the interpretation is the same, viz. that on the margin, higher types have a greater willingness to pay for increasing the outcome $y \in \mathcal{Y}$. It is worth noting, however, that whereas the classical Spence–Mirrlees condition is (nearly) ordinal,[20] the outer Spence–Mirrlees condition is not.

4.3 Increasing allocations are implementable

[16] The continuity of $f_3(y, \cdot, t)$ plays a technical role in the proof: see footnote 21 below.

[17] Given $\mathcal{T} \subseteq \mathbf{R}$, a function $\phi : \mathcal{T} \to \mathbf{R}$ is called *single-crossing* iff for any $t < t'$ in $\mathcal{T}$, $\phi(t) \geq (>) \, 0$ implies $\phi(t') \geq (>) \, 0$, and *strictly single-crossing* iff $\phi(t) \geq 0$ implies $\phi(t') > 0$.

[18] An equivalent definition of the Spence–Mirrlees condition requires instead that the slope $f_1(y,p,t)/|f_2(y,p,t)|$ of the agent's indifference curve through any point $(y,p) \in \mathcal{Y} \times \mathbf{R}$ be increasing in t. See Milgrom and Shannon (1994, Theorem 3) for a proof of equivalence.

[19] The *upper derivative* of $\phi : [0,1] \to \mathbf{R}$ at $t \in (0,1)$ is $\frac{\overline{\mathrm{d}}}{\overline{\mathrm{d}}m}\phi(t+m)\big|_{m=0} \coloneqq \limsup_{m \to 0} \left[\phi(t+m) - \phi(t)\right]/m$. Nothing changes in the sequel if the upper derivative is replaced with the lower (defined with a lim inf), or with any of the four Dini derivatives.

[20] Precisely: if f satisfies this condition, then so does $\phi \circ f$ for any differentiable and strictly increasing transformation $\phi : \mathbf{R} \to \mathbf{R}$.

Implementability theorem. If $\mathcal{Y}$ and f are regular and f satisfies the outer Spence–Mirrlees condition, then any increasing allocation is implementable.

The proof is in appendix H. The idea is as follows. Take any increasing allocation $Y : [0,1] \to \mathcal{Y}$. By the existence lemma in appendix H.1, there exists a payment schedule $P : [0,1] \to \mathbf{R}$ such that (Y, P) satisfies the envelope formula.[21] By the converse envelope theorem, it follows that (Y, P) is locally incentive-compatible in the sense that it satisfies the outer first-order condition. The outer Spence–Mirrlees condition ensures that local incentive-compatibility translates into global incentive-compatibility.

The argument for the final step actually applies only to allocations Y that are suitably continuous. But the regularity of $\mathcal{Y}$ ensures (via a lemma in appendix H.2) that any increasing Y can be approximated by a sequence of continuous and increasing (hence implementable) allocations.

Given two mild additional assumptions, the payment rule implementing a given increasing allocation is in fact unique, and may be computed constructively via Picard's method—see appendix H.1.

The implementability theorem admits a standard converse when $\mathcal{Y}$ is a chain (e.g. an interval of $\mathbf{R}$), proved in appendix I:

Proposition 1. If $\mathcal{Y}$ and f are regular, f satisfies the strict outer Spence–Mirrlees condition, and $\mathcal{Y}$ is a chain, then all and only increasing allocations are implementable.

4.4 Selling information

In this section, I apply the implementability theorem to selling informative signals. Here the outcomes $\mathcal{Y}$ are distributions of posterior beliefs—a space very different from an interval of $\mathbf{R}$. I show that all Blackwell-increasing information allocations are implementable, and that only these are implementable if agents are able to share information with each other.

There is a population of agents with types $t \in [0,1]$, a finite set Ω of states of the world, and a set A of actions. A type-t agent earns payoff $U(a, \omega, t)$ if she takes action $a \in A$ in state $\omega \in \Omega$, so her expected value at belief $\mu \in \Delta(\Omega)$ is

$$V(\mu, t) := \sup_{a \in A} \sum_{\omega \in \Omega} U(a, \omega, t)\mu(\omega).$$

Assume that the type derivative V_2 exists and is bounded, and that $V_2(\cdot, t)$ is continuous for each $t \in [0,1]$.[22]

[21] This is where the continuity of $f_3(y, \cdot, t)$ is used: the existence lemma requires it.

[22] This is slightly stronger than assuming that the underlying type derivative U_3 has the same

Example 2. Each agent is tasked with announcing a probabilistic forecast $a \in A \coloneqq \Delta(\Omega)$ of the state $\omega \in \Omega$. Ex post, the public's assessment of an agent's quality as a forecaster is some function of the forecast a and realised state ω (a *scoring rule*); for concreteness, $a(\omega)/\|a\|_2$, where $\|\cdot\|_2$ denotes the Euclidean norm.[23] Each agent attaches some importance $t \in [0,1]$ to being considered a good forecaster, so that $U(a, \omega, t) = ta(\omega)/\|a\|_2$. Agents are expected-utility maximisers.

It is easily verified that an agent with belief $\mu \in \Delta(\Omega)$ optimally announces forecast $a = \mu$. Her value is therefore

$$V(\mu, t) = \sum_{\omega \in \Omega} \frac{t\mu(\omega)}{\|\mu\|_2} \mu(\omega) = t\|\mu\|_2.$$

By inspection, $V_2(\mu, t) = \|\mu\|_2$ exists, is bounded, and is continuous in μ. ◇

Agents share a common prior $\mu_0 \in \operatorname{int} \Delta(\Omega)$. Before making her decision, an agent observes the realisation of a signal (a random variable correlated with ω), and forms a posterior belief according to Bayes's rule. Since the signal is random, the agent's posterior is random; write y for its distribution (a Borel probability measure on $\Delta(\Omega)$). The agent's expected payoff under a signal that induces posterior distribution y, if she makes payment $p \in \mathbf{R}$, is

$$f(y, p, t) \coloneqq g\left(\int_{\Delta(\Omega)} V(\mu, t) y(\mathrm{d}\mu), p\right),$$

where $g : \mathbf{R}^2 \to \mathbf{R}$ is jointly continuous, possesses a bounded derivative g_1 that is continuous in p, and has $g(v, \cdot)$ strictly decreasing and onto $\mathbf{R}$ for each $v \in \mathbf{R}$. The payoff f is regular: f_3 exists, is bounded, and is continuous in p, and I verify the joint continuity property in supplemental appendix N.

A Borel probability measure y on $\Delta(\Omega)$ is the distribution of posteriors induced by some signal exactly if its mean $\int_{\Delta(\Omega)} \mu y(\mathrm{d}\mu)$ is equal to μ_0.[24] Write $\mathcal{Y}$ for the set of all mean-μ_0 distributions of posteriors, and order it by Blackwell informativeness: $y \lesssim y'$ iff $\int_{\Delta(\Omega)} v \mathrm{d}y \leq \int_{\Delta(\Omega)} v \mathrm{d}y'$ for every continuous and convex $v : \Delta(\Omega) \to \mathbf{R}$.[25]

properties; see e.g. Milgrom and Segal (2002, Theorem 3) for sufficient conditions.

[23] More generally, any bounded and strictly proper scoring rule will do. See e.g. Gneiting and Raftery (2007) for an introduction to proper scoring rules.

[24] The 'only if' direction is trivial. Conversely, a y with mean μ_0 is induced by a $\Delta(\Omega)$-valued signal whose distribution conditional on each $\omega \in \Omega$ is

$$\pi(M|\omega) = \frac{1}{\mu_0(\omega)} \int_M \mu(\omega) y(\mathrm{d}\mu) \quad \text{for each Borel-measurable } M \subseteq \Delta(\Omega).$$

This construction is due to Blackwell (1951), and used by Kamenica and Gentzkow (2011).

[25] A Blackwell-less informative distribution of posteriors is precisely one that yields a lower expected payoff $\int_{\Delta(\Omega)} V(\mu, t) y(\mathrm{d}\mu)$ no matter what the underlying action set A or utility $U(\cdot, \cdot, t)$. This is because $V(\cdot, t)$ is continuous and convex for any A and U, and any continuous and convex v can be approximated by $V(\cdot, t)$ for some A and U.

I show in supplemental appendix L that the outcome space $\mathcal{Y}$ is regular.

Assume that f satisfies the strict outer Spence–Mirrlees condition. An *information allocation* is a map $Y : [0,1] \to \mathcal{Y}$ that assigns to each type a distribution of posteriors. By the implementability theorem, we have:

Proposition 2. Every increasing information allocation is implementable.

The converse is false. In particular, there are implementable allocations that assign some types $t < t'$ Blackwell-incomparable information. But any such information allocation is vulnerable to collusion, as agents of types t and t' would benefit by sharing their information.[26,27] Call an allocation *sharing-proof* iff no two types are assigned Blackwell-incomparable information.

Proposition 3. An information allocation is implementable and sharing-proof if and only if it is increasing.

The proof is in appendix J.

Appendix to the theory (§2 and §3)

A Mathematical background

Two operations are important in this paper: writing a function as the integral of its derivative, and interchanging limits and integrals. The former is permissible precisely for absolutely continuous functions:

Definition 8. A function $\phi : [0,1] \to \mathbf{R}$ is *absolutely continuous* iff for each $\varepsilon > 0$, there is a $\delta > 0$ such that for any finite collection $\{(r_n, t_n)\}_{n=1}^N$ of disjoint intervals of $[0,1]$, $\sum_{n=1}^N (t_n - r_n) < \delta$ implies $\sum_{n=1}^N |\phi(t_n) - \phi(r_n)| < \varepsilon$.

Absolute continuity implies continuity and differentiability a.e., but the converse is false. Absolute continuity is implied by Lipschitz continuity.

Lebesgue's fundamental theorem of calculus.[28] Let ϕ be a function $[0,1] \to \mathbf{R}$. The following are equivalent:

(1) ϕ is absolutely continuous.

[26] This holds no matter how the underlying signals giving rise to the posterior distributions $Y(t)$ and $Y(t')$ are correlated with each other. For by a standard embedding theorem (e.g. Theorem 7.A.1 in Shaked and Shanthikumar (2007)), $Y(t) \lesssim Y(t')$ is necessary (as well as sufficient) for there to exist a probability space on which there are random vectors with laws $Y(t)$ and $Y(t')$ such that the latter is statistically sufficient for the former.

[27] Both agents benefit *strictly* provided $V(\cdot, t)$ and $V(\cdot, t')$ are strictly convex.

[28] See e.g. Folland (1999, §3.5, p. 106) for a proof.

(2) There is a $\psi \in \mathcal{L}^1$ such that $\phi(t) = \phi(0) + \int_0^t \psi$ for every $t \in [0,1]$.

(3) ϕ is differentiable a.e., its (a.e.-defined) derivative ϕ' belongs to $\mathcal{L}^1$, and $\phi(t) = \phi(0) + \int_0^t \phi'$ for every $t \in [0,1]$.

As for interchanging limits and integrals, uniform integrability is the key:

Definition 9. A family $\Phi \subseteq \mathcal{L}^1$ is *uniformly integrable* iff for each $\varepsilon > 0$, there is $\delta > 0$ such that for any open $T \subseteq [0,1]$ of measure $< \delta$, we have $\int_T |\phi| < \varepsilon$ for every $\phi \in \Phi$.

Vitali convergence theorem.[29] Let $\{\phi_n\}_{n \in \mathbf{N}}$ be a uniformly integrable sequence in $\mathcal{L}^1$ converging a.e. to $\phi : [0,1] \to \mathbf{R}$. Then $\phi \in \mathcal{L}^1$, and $\lim_{n\to\infty} \int_r^t \phi_n = \int_r^t \phi$ for all $r, t \in [0,1]$.

(Lebesgue's dominated convergence theorem is a corollary.)

Absolute continuity and uniform integrability are closely related:

AC–UI lemma (Fitzpatrick & Hunt, 2015)**.** Let ϕ be a continuous function $[0,1] \to \mathbf{R}$. The following are equivalent:

(1) ϕ is absolutely continuous.

(2) The 'divided-difference' family $\{t \mapsto [\phi(t+m) - \phi(t)]/m\}_{m>0}$ is uniformly integrable.

B Housekeeping for absolute equi-continuity (§2.1, p. 4)

The following lemma justifies the name 'absolute equi-continuity', and is used in appendix E below to prove the necessity lemma (§3.1, p. 8).

Lemma 1. An absolutely equi-continuous family $\{\phi_x\}_{x \in \mathcal{X}}$ is uniformly equi-continuous, and each of its members ϕ_x is absolutely continuous.

Proof. Let $\{\phi_x\}_{x \in \mathcal{X}}$ be absolutely equi-continuous. Then for every $x \in \mathcal{X}$, $\{t \mapsto [\phi_x(t+m) - \phi_x(t)]/m\}_{m>0}$ is uniformly integrable, and hence ϕ_x is absolutely continuous by the AC–UI lemma in appendix A.

It follows that for any $r < t$ in $[0,1]$,

$$\sup_{x\in\mathcal{X}} |\phi_x(t) - \phi_x(r)| = \sup_{x\in\mathcal{X}} \left| \int_r^t \phi_x' \right| = \sup_{x\in\mathcal{X}} \left| \lim_{m\downarrow 0} \int_r^t \frac{\phi_x(s+m) - \phi_x(s)}{m} \mathrm{d}s \right|$$
$$\leq \sup_{x\in\mathcal{X}} \sup_{m>0} \left| \int_r^t \frac{\phi_x(s+m) - \phi_x(s)}{m} \mathrm{d}s \right| \leq \sup_{m>0} \int_r^t \sup_{x\in\mathcal{X}} \left| \frac{\phi_x(s+m) - \phi_x(s)}{m} \right| \mathrm{d}s,$$

[29] For a proof and a partial converse, see e.g. Royden and Fitzpatrick (2010, §4.6).

where the first equality holds by Lebesgue's fundamental theorem of calculus, and the second holds by the Vitali convergence theorem.

Fix an $\varepsilon > 0$. By the absolute equi-continuity of $\{\phi_x\}_{x\in\mathcal{X}}$, there is a $\delta > 0$ such that whenever $t - r < \delta$, the right-hand side of the above inequality is $< \varepsilon$, and thus $\sup_{x\in\mathcal{X}}|\phi_x(t) - \phi_x(r)| < \varepsilon$. So $\{\phi_x\}_{x\in\mathcal{X}}$ is uniformly equi-continuous. ■

C Proof of the identity lemma (§3.2, p. 9)

We use the results in appendix A. We shall focus on the limit $m \downarrow 0$, omitting the symmetric argument for $m \uparrow 0$.[30] For $t \in [0, 1)$ and $m \in (0, 1 - t]$, write

$$\begin{aligned}\phi_m(t) &\coloneqq \frac{V_X(t+m) - V_X(t)}{m} \\ &= \underbrace{\frac{f(X(t+m), t+m) - f(X(t+m), t)}{m}}_{=:\, \psi_m(t)} + \underbrace{\frac{f(X(t+m), t) - f(X(t), t)}{m}}_{=:\, \chi_m(t)}.\end{aligned}$$

Fix $r, t \in (0, 1)$. Note that

$$\lim_{m\downarrow 0} \int_r^t \chi_m = \frac{\mathrm{d}}{\mathrm{d}m} \int_r^t f(X(s+m), s)\mathrm{d}s \bigg|_{m=0}$$

whenever the limit exists. Our task is to show that $\{\int_r^t \chi_m\}_{m>0}$ is convergent as $m \downarrow 0$ with limit

$$V_X(t) - V_X(r) - \int_r^t f_2(X(s), s)\mathrm{d}s.$$

$\{\psi_m\}_{m>0}$ need not converge a.e. under the basic assumptions.[31] But

$$\psi^\star_m(t) \coloneqq \frac{f(X(t), t) - f(X(t), t-m)}{m}$$

converges pointwise to $t \mapsto f_2(X(t), t)$, and by a change of variable,

$$\int_r^t \psi_m = \int_{r+m}^{t+m} \psi^\star_m = \int_r^t \psi^\star_m + \left(\int_t^{t+m} \psi^\star_m - \int_r^{r+m} \psi^\star_m \right) = \int_r^t \psi^\star_m + \mathrm{o}(1),$$

where the bracketed terms vanish as $m \downarrow 0$ because $\{\psi^\star_m\}_{m>0}$ is uniformly integrable by the basic assumptions.

By absolute continuity of V_X and the AC–UI lemma in appendix A, $\{\phi_m\}_{m>0}$ is

[30] Since the argument below relies on absolute equi-continuity, the omitted argument requires uniform integrability of $\{\Phi_m\}_{m<0} \coloneqq \{t \mapsto \sup_{x\in\mathcal{X}}|(f(x, t+m) - f(x, t))/m|\}_{m<0}$. This follows from absolute equi-continuity and the observation that $\Phi_m(t) = \Phi_{-m}(t + m)$.

[31] This remains true even under much stronger assumptions. For example, equi-differentiability of $\{f(x, \cdot)\}_{x\in\mathcal{X}}$ is not enough: a counter-example is $\mathcal{X} = [0, 1]$, $f(x, t) = (t - x)\mathbf{1}_{\mathbf{Q}}(x)$ and $X(t) = t$. (Here $\mathbf{1}_{\mathbf{Q}}(x) = 1$ if x is rational and $= 0$ otherwise.) In this case $\psi_m(t) = \mathbf{1}_{\mathbf{Q}}(t + m)$, which is nowhere convergent as $m \downarrow 0$.

uniformly integrable and converges a.e. to V_X' as $m \downarrow 0$. Since $\{\psi_m^\star\}_{m>0}$ is uniformly integrable and converges pointwise to $t \mapsto f_2(X(t), t)$, it follows that

$$\begin{aligned}\lim_{m\downarrow 0}\int_r^t \chi_m = \lim_{m\downarrow 0}\int_r^t [\phi_m - \psi_m] &= \lim_{m\downarrow 0}\int_r^t [\phi_m - \psi_m^\star] \\ &= \int_r^t \lim_{m\downarrow 0}[\phi_m - \psi_m^\star] = \int_r^t [V_X'(s) - f_2(X(s), s)]\,\mathrm{d}s,\end{aligned}$$

where the third equality holds by the Vitali convergence theorem. Since the last expression is well-defined, this shows $\{\int_r^t \chi_m\}_{m>0}$ to be convergent as $m \downarrow 0$. And because V_X is absolutely continuous, the value of the limit is

$$\lim_{m\downarrow 0}\int_r^t \chi_m = V_X(t) - V_X(r) - \int_r^t f_2(X(s), s)\mathrm{d}s$$

by Lebesgue's fundamental theorem of calculus. ■

D A characterisation of absolute continuity of the value

The following lemma characterises the absolute-continuity-of-V_X condition that appears in the main theorem (§3.2, p. 8). Apart from its independent interest, it is needed for the proofs in appendices E and F below.

Lemma 2. Under the basic assumptions, the following are equivalent:

(1) $V_X(t) \coloneqq f(X(t), t)$ is absolutely continuous.

(2) The family $\{\chi_m\}_{m>0}$ is uniformly integrable, where

$$\chi_m(t) \coloneqq \frac{f(X(t+m), t) - f(X(t), t)}{m}.$$

In the classical case, (2) is imposed (it follows from the classical assumptions, by Lemma 4 in appendix F below). In the modern case, (1) arises within the theorem. Both are clearly joint restrictions on f and X.[32]

Proof. Define $\{\phi_m\}_{m>0}$ and $\{\psi_m\}_{m>0}$ as in the proof of the identity lemma (appendix C). $\{\psi_m\}_{m>0}$ is uniformly integrable by the basic assumption of absolute equi-continuity. By the AC–UI lemma in appendix A, (1) is equivalent to $\{\phi_m\}_{m>0}$ being uniformly integrable.

Suppose that $\{\chi_m\}_{m>0}$ is uniformly integrable, and fix $\varepsilon > 0$. Let $\delta > 0$ meet the $\varepsilon/2$-challenge for both $\{\psi_m\}_{m>0}$ and $\{\chi_m\}_{m>0}$; then for any open $T \subseteq [0, 1]$ of

[32] As emphasised by Milgrom and Segal (2002), however, any *optimal* X satisfies (1) provided f satisfies the basic assumptions. See appendix E below for a proof.

measure $< \delta$ and any $m > 0$, we have

$$\int_T |\phi_m| \leq \int_T |\psi_m| + \int_T |\chi_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

showing that $\{\phi_m\}_{m>0}$ is uniformly integrable.

An almost identical argument establishes that uniform integrability of $\{\phi_m\}_{m>0}$ implies uniform integrability of $\{\chi_m\}_{m>0}$. ■

E Proof of the necessity lemma (§3.1, p. 8)

Lemma 3. If $\{f(x,\cdot)\}_{x\in\mathcal{X}}$ is absolutely equi-continuous, then the value $V_X(t) \coloneqq f(X(t),t)$ of any optimal $X : [0,1] \to \mathcal{X}$ is absolutely continuous.

Proof. Let X be optimal. Then for any $r < t$ in $[0,1)$ and $m \in (0, 1-t]$,

$$\begin{aligned} \left|\frac{1}{m}\int_t^{t+m} V_X - \frac{1}{m}\int_r^{r+m} V_X\right| &= \left|\int_r^t \frac{V_X(s+m) - V_X(s)}{m}\mathrm{d}s\right| \\ &\leq \int_r^t \left|\frac{V_X(s+m) - V_X(s)}{m}\right|\mathrm{d}s \leq \int_r^t D_m, \end{aligned}$$

where

$$D_m(s) \coloneqq \sup_{x\in\mathcal{X}}\left|\frac{f(x,s+m) - f(x,s)}{m}\right|.$$

Fix an $\varepsilon > 0$. The absolute equi-continuity of $\{f(x,\cdot)\}_{x\in\mathcal{X}}$ provides that $\{D_m\}_{m>0}$ is uniformly integrable, so that there is a $\delta > 0$ such that for any open $T \subseteq [0,1]$ of measure $< \delta$, we have $\int_T D_m < \varepsilon/2$ for every $m > 0$. Thus for any finite collection $\{(r_n,t_n)\}_{n=1}^N$ of disjoint open intervals of $[0,1]$ whose union T has measure $< \delta$, we have

$$\sum_{n=1}^N \left|\frac{1}{m}\int_{t_n}^{t_n+m} V_X - \frac{1}{m}\int_{r_n}^{r_n+m} V_X\right| \leq \int_T D_m < \varepsilon/2 \quad \text{for every } m > 0.$$

V_X is (uniformly) continuous since $\{f(x,\cdot)\}_{x\in\mathcal{X}}$ is uniformly equi-continuous by Lemma 1 in appendix B.[33] Thus letting $m \downarrow 0$ yields

$$\sum_{n=1}^N |V_X(t_n) - V_X(r_n)| \leq \varepsilon/2 < \varepsilon$$

by the mean-value theorem, showing V_X to be absolutely continuous. ■

Proof of the necessity lemma. Let X be optimal, and fix $r < t$ in $[0,1]$. V_X is

[33] For any $\varepsilon > 0$, the uniform equi-continuity of $\{f(x,\cdot)\}_{x\in\mathcal{X}}$ delivers a $\delta > 0$ such that $|t-r| < \delta$ implies $|V_X(t) - V_X(r)| \leq \sup_{x\in\mathcal{X}}|f(x,t) - f(x,r)| < \varepsilon$.

absolutely continuous by Lemma 3. Define $\phi_{r,t} : [-r, 1-t] \to \mathbf{R}$ by

$$\phi_{r,t}(m) \coloneqq \int_r^t f(X(s+m), s)\mathrm{d}s \quad \text{for each } m \in [-r, 1-t].$$[34]

$\phi'_{r,t}(0)$ exists by the identity lemma (§3.2, p. 9). To show that it is zero, observe that for any $s \in (r,t)$ and $m \in (0, \min\{s, 1-s\}]$, optimality requires

$$\frac{f(X(s+m), s) - f(X(s), s)}{m} \leq 0 \leq \frac{f(X(s-m), s) - f(X(s), s)}{-m}.$$

Integrating over (r,t) and letting $m \downarrow 0$ yields $\phi'_{r,t}(0) \leq 0 \leq \phi'_{r,t}(0)$. ■

F A lemma under the classical assumptions

The following result is used in the proof of the housekeeping lemma (§3.1, p. 7), as well as in the proof of the classical envelope theorem and converse in appendix G below.

Lemma 4. Fix a decision rule $X : [0,1] \to \mathcal{X}$, and let

$$\chi_m(t) \coloneqq \frac{f(X(t+m), t) - f(X(t), t)}{m}.$$

(1) Under the basic and classical assumptions, $\{\chi_m\}_{m>0}$ is uniformly integrable.

(2) Under the basic assumptions, the following are equivalent:

(a) $\{\chi_m\}_{m>0}$ is uniformly integrable and convergent a.e. as $m \downarrow 0$.

(b) $V_X(t) \coloneqq f(X(t), t)$ is absolutely continuous, and the derivative $\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m), t)\Big|_{m=0}$ exists for a.e. $t \in (0,1)$.

Proof. For (1), write K for the vector of non-negative constants that bounds f_1, and $L \geq 0$ for the Lipschitz constant of X. Let $\|\cdot\|_2$ denote the Euclidean norm. For any $t \in [0,1)$ and $m \in (0, 1-t]$, writing $x_\omega \coloneqq (1-\omega)X(t) + \omega X(t+m)$ for $\omega \in [0,1]$, we have by the Cauchy–Schwarz inequality that

$$\begin{aligned} |\chi_m(t)| &= \left| \frac{1}{m} \int_0^1 \Big(f_1(x_\omega, t) \cdot [X(t+m) - X(t)] \Big) \mathrm{d}\omega \right| \\ &\leq \frac{1}{m} \int_0^1 \Big(\|f_1(x_\omega, t)\|_2 \times \|X(t+m) - X(t)\|_2 \Big) \mathrm{d}\omega \leq \frac{1}{m} \|K\|_2 \times Lm = \|K\|_2 L. \end{aligned}$$

Thus $\{\chi_m\}_{m>0}$ is uniformly bounded, hence uniformly integrable.

[34]The map $s \mapsto f(X(s+m), s)$ is integrable because $|f(X(s+m), s)| \leq |V_X(s)| + |f(X(s+m), s) - f(X(s), s)|$, where the former term is continuous, and the latter is integrable by Lemma 2 in appendix D.

For (2), absolute continuity of V_X is equivalent to uniform integrability of $\{\chi_m\}_{m>0}$ by Lemma 2 in appendix D, and a.e. existence of $\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)|_{m=0}$ is definitionally equivalent to a.e. convergence of $\{\chi_m\}_{m>0}$. ■

G Proof of the classical envelope theorem and converse (§2.2)

Proof. Fix a Lipschitz continuous decision rule $X : [0,1] \to \mathcal{X}$. By Lemma 4 in appendix F, $V_X(t) \coloneqq f(X(t),t)$ is absolutely continuous, hence differentiable a.e. The map $r \mapsto f(X(r),t)$ is differentiable a.e. by the classical assumptions, and $t \mapsto f(X(r),t)$ is differentiable by the basic assumptions. Hence the a.e.-defined derivative of V_X obeys the differentiation identity

$$V_X'(t) = \frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)\bigg|_{m=0} + f_2(X(t),t) \quad \text{for a.e. } t \in (0,1).$$

It follows that the first-order condition a.e. is equivalent to

$$V_X'(t) = f_2(X(t),t) \quad \text{for a.e. } t \in (0,1),$$

which in turn is equivalent to the envelope formula by Lebesgue's fundamental theorem of calculus. ■

By inspection, the proof requires precisely absolute continuity of V_X (so that the envelope formula can be satisfied) and a.e. existence of $\frac{\mathrm{d}}{\mathrm{d}m} f(X(t+m),t)|_{m=0}$ (so that the first-order condition a.e. is well-defined). Part (2) of Lemma 4 in appendix F therefore tells us that the classical assumptions can be weakened to uniform integrability and a.e. convergence of $\{\chi_m\}_{m>0}$, and no further. For f non-trivial, the uniform integrability part involves a strong continuity requirement on X.[35]

Appendix to the application (§4)

H Proof of the implementability theorem (§4.3, p. 13)

We state two lemmata in §H.1–§H.2, then prove the theorem in §H.3.

[35] For example, consider $\mathcal{X} = [0,1]$, $f(x,t) = x$ and $X(t) = \mathbf{1}_{[r,1]}$, where $r \in (0,1)$. Then given $m > 0$, we have $\chi_m(t) = 1/m$ for all $t \in [r-m, r]$. Suppose toward a contradiction that $\{\chi_m\}_{m>0}$ is uniformly integrable, and let $\delta > 0$ meet the ε-challenge for $\varepsilon \in (0,1)$; then for all $m \in (0, \delta/2)$, we have $\int_{r-\delta/2}^{r+\delta/2} |\chi_m| \geq \int_{r-m}^{r} |\chi_m| = m/m = 1 > \varepsilon$, which is absurd. This example clearly generalises: the gist is that uniform integrability of $\{\chi_m\}_{m>0}$ is incompatible with non-removable discontinuities in X unless f is trivial.

H.1 Solutions of the envelope formula

In the first step of the argument in §H.3 below, we are given an allocation Y, and wish to choose a payment schedule P such that (Y, P) satisfies the envelope formula. The following asserts that this can be done:

Existence lemma. Assume that for all $(y,t) \in \mathcal{Y} \times [0,1]$, $f(y,\cdot,t)$ is strictly decreasing, continuous and onto $\mathbf{R}$. Further assume that the type derivative f_3 exists and is bounded, and that $f_3(y,\cdot,t)$ is continuous for all $(y,t) \in \mathcal{Y} \times [0,1]$. Then for any $k \in \mathbf{R}$ and any allocation $Y : [0,1] \to \mathcal{Y}$ such that $t \mapsto f(Y(t),p,t)$ and $t \mapsto f_3(Y(t),p,t)$ are Borel-measurable for every $p \in \mathbf{R}$, there exists a payment schedule $P : [0,1] \to \mathbf{R}$ such that (Y,P) satisfies the envelope formula with $V_{Y,P}(0) = k$.

Remark 2. The following corollary may prove useful elsewhere: suppose in addition that $\mathcal{Y}$ is equipped with some topology such that $f(\cdot,p,t)$ and $f_3(\cdot,p,t)$ are Borel-measurable and $f_3(y,p,\cdot)$ is continuous. Then for any Borel-measurable allocation $Y : [0,1] \to \mathcal{Y}$, there is a payment schedule P such that (Y,P) satisfies the envelope formula.

The existence lemma is immediate from the following abstract result by letting $\phi(p,t) \coloneqq f(Y(t),p,t)$ and $\psi(p,t) \coloneqq f_3(Y(t),p,t)$.

Lemma 5. Let ϕ and ψ be functions $\mathbf{R} \times [0,1] \to \mathbf{R}$. Suppose that $\phi(\cdot,t)$ is strictly decreasing, continuous, and onto $\mathbf{R}$ for every $t \in [0,1]$, and that ψ is bounded with $\psi(\cdot,t)$ continuous for every $t \in [0,1]$. Further assume that $\phi(p,\cdot)$ and $\psi(p,\cdot)$ are Borel-measurable for each $p \in \mathbf{R}$. Then for any $k \in \mathbf{R}$, there is a function $P : [0,1] \to \mathbf{R}$ such that

$$\phi(P(t),t) = k + \int_0^t \psi(P(s),s)\mathrm{d}s \quad \text{for every } t \in [0,1].$$

Proof. Since $\phi(\cdot,t)$ is strictly decreasing and continuous, it possesses a continuous inverse $\phi^{-1}(\cdot,t)$, well-defined on all of $\mathbf{R}$ since $\phi(\mathbf{R},t) = \mathbf{R}$. We may therefore define a function $\chi : \mathbf{R} \times [0,1] \to \mathbf{R}$ by

$$\chi(w,t) \coloneqq \psi\left(\phi^{-1}(w,t),t\right) \quad \text{for each } w \in \mathbf{R} \text{ and } t \in [0,1].$$

$\chi(\cdot,t)$ is continuous since $\psi(\cdot,t)$ and $\phi^{-1}(\cdot,t)$ are, χ is bounded since ψ is, and $\chi(w,\cdot)$ is Borel-measurable since $\psi(\cdot,t)$ is continuous and $\psi(p,\cdot)$ and $\phi^{-1}(w,\cdot)$ are Borel-measurable.

Fix $k \in \mathbf{R}$. Consider the integral equation

$$W(t) = k + \int_0^t \chi(W(s),s)\mathrm{d}s \quad \text{for } t \in [0,1],$$

where W is an unknown function $[0,1] \to \mathbf{R}$. Since $\chi(\cdot, t)$ is continuous and $\chi(w, \cdot)$ bounded and Borel-measurable, there is a local solution by Carathéodory's existence theorem;[36] call it V. By boundedness of χ and a comparison theorem,[37] V can be extended to a solution on all of $[0,1]$.

Now define $P(t) \coloneqq \phi^{-1}(V(t), t)$. For every $t \in [0,1]$, it satisfies

$$\phi(P(t), t) = V(t) = k + \int_0^t \chi(V(s), s) \mathrm{d}s = k + \int_0^t \psi(P(s), s) \mathrm{d}s. \qquad \blacksquare$$

Uniqueness corollary. Under the hypotheses of the existence lemma, if in addition $\{f_3(y, \cdot, t)\}_{(y,t) \in \mathcal{Y} \times [0,1]}$ is Lipschitz equi-continuous[38] and the monotonicity of $f(y, \cdot, t)$ is uniform in the sense that for some $M > 0$,

$$f(y, p, t) - f(y, p', t) \geq M(p' - p) \quad \text{for any } p < p' \text{ in } \mathbf{R},\ y \in \mathcal{Y} \text{ and } t \in [0,1],$$

then there is *exactly one* payment schedule P such that (Y, P) satisfies the envelope formula with $V_{Y,P}(0) = k$, and this payment schedule may be computed via Picard's method.

Proof. Again let $\phi(p, t) \coloneqq f(Y(t), p, t)$ and $\psi(p, t) \coloneqq f_3(Y(t), p, t)$, and return to the proof of Lemma 5. The additional assumptions ensure, respectively, that $\{\psi(\cdot, t)\}_{t \in [0,1]}$ and $\{\phi^{-1}(\cdot, t)\}_{t \in [0,1]}$ are Lipschitz equi-continuous. In follows that $\{\chi(\cdot, t)\}_{t \in [0,1]}$ is Lipschitz equi-continuous, so that (the Picard operator is a contraction, and thus) the integral equation has a unique solution to which Picard iteration converges in the sup norm.[39] ■

H.2 Continuous approximation of increasing maps

The second step of the argument §H.3 below relies on approximating an increasing map $[0,1] \to \mathcal{Y}$ by continuous and increasing maps. This is made possible by the following:

Approximation lemma. Let $\mathcal{Y}$ be regular, and let Y be an increasing map $[0,1] \to \mathcal{Y}$. The image $Y([0,1])$ may be embedded in a chain $\mathcal{C} \subseteq \mathcal{Y}$ with $\inf \mathcal{C} = Y(0)$ and $\sup \mathcal{C} = Y(1)$ that is order-dense-in-itself, order-complete and order-separable.[40] Furthermore, there exists a sequence $(Y_n)_{n \in \mathbf{N}}$ of increasing maps $[0,1] \to \mathcal{C}$, each with $Y_n = Y$ on $\{0,1\}$, such that when $\mathcal{C}$ has the relative topology inherited from

[36] See e.g. Theorem 5.1 in Hale (1980, ch. 1).

[37] See e.g. Theorem 2.17 in Teschl (2012).

[38] That is, there is an $L \geq 0$ such that $f_3(y, \cdot, t)$ is L-Lipschitz for every $(y,t) \in \mathcal{Y} \times [0,1]$.

[39] See e.g. Theorem 5.3 in Hale (1980, ch. 1).

[40] $\mathcal{C} \subseteq \mathcal{Y}$ is *order-complete* iff every subset with a lower (upper) bound has an infimum (supremum), and *order-separable* iff it has a countable order-dense subset.

the order topology on $\mathcal{Y}$, Y_n is continuous for each $n \in \mathbf{N}$, and $Y_n \to Y$ pointwise as $n \to \infty$.

The (rather involved) proof is in supplemental appendix M.

H.3 Proof of the implementability theorem

Fix an increasing $Y : [0,1] \to \mathcal{Y}$. Embed its image $Y([0,1])$ in the chain $\mathcal{C} \subseteq \mathcal{Y}$ delivered by the approximation lemma in appendix H.2, and equip $\mathcal{C}$ with the relative topology inherited from the order topology on $\mathcal{Y}$. We henceforth view Y as a function $[0,1] \to \mathcal{C}$, and (with a minor abuse of notation) view f and f_3 as functions $\mathcal{C} \times \mathbf{R} \times [0,1] \to \mathbf{R}$.

We seek a payment schedule $P : [0,1] \to \mathbf{R}$ such that the direct mechanism (Y, P) is incentive-compatible. We do this first (step 1) under the assumption that Y is continuous, then (step 2) show how continuity may be dropped.

Step 1: Suppose that Y is continuous. By preference regularity and the existence lemma in appendix H.1,[41] there exists a payment schedule $P : [0,1] \to \mathbf{R}$ such that the envelope formula holds with (say) $V_{Y,P}(0) = 0$:

$$V_{Y,P}(t) = \int_0^t f_3(Y(s), P(s), s)\mathrm{d}s \quad \text{for every } t \in [0,1].$$

This P must be continuous since Y, f and $V_{Y,P}$ are continuous and $f(y, \cdot, t)$ is strictly monotone.[42] We will show that (Y, P) is incentive-compatible.

Write $U(r,t) \coloneqq f(Y(r), P(r), t)$ for type t's mimicking payoff, and $\phi_{r,t}(m) \coloneqq \int_r^t U(s+m, s)\mathrm{d}s$ for the collective payoff of types $[r,t] \subseteq (0,1)$ from 'mimicking up' by m. Clearly U is a continuous function $[0,1]^2 \to \mathbf{R}$, and thus $\phi_{r,t} : [-r, 1-t] \to \mathbf{R}$ is also continuous. Note that $V_{Y,P}(t) \equiv U(t,t)$.

The model fits into the abstract setting of §2.1 by letting $\mathcal{X} \coloneqq \mathcal{C} \times \mathbf{R}$ and $X(t) \coloneqq (Y(t), P(t))$, and the basic assumptions are satisfied since f_3 exists and is bounded. We may thus invoke the converse envelope theorem (p. 8): since (Y, P)

[41] The measurability hypothesis in the existence lemma is satisfied because $f(\cdot, p, t)$, $f_3(\cdot, p, t)$ and Y are continuous, and $f(y, p, \cdot)$ and $f_3(y, p, \cdot)$ are Borel-measurable (the former being continuous, and the latter a derivative). (To complete the argument for measurability, deduce that $r \mapsto f(Y(r), p, t)$ is continuous and that $t \mapsto f(Y(r), p, t)$ is Borel-measurable, so that $(r,t) \mapsto f(Y(r), p, t)$ is (jointly) Borel-measurable, and thus $t \mapsto f(Y(t), p, t)$ is Borel-measurable. Similarly for f_3.)

[42] Suppose not: $t_n \to t$ but $\lim_{n\to\infty} P(t_n) \neq P(t)$. Then the continuity of Y and f and the strict monotonicity of $f(y, \cdot, t)$ yield a contradiction with the continuity of $V_{Y,P}$:

$$V_{Y,P}(t_n) = f(Y(t_n), P(t_n), t_n) \to f\left(Y(t), \lim_{n\to\infty} P(t_n), t\right) \neq f(Y(t), P(t), t) = V_{Y,P}(t).$$

satisfies the envelope formula, it must satisfy the outer first-order condition:

$$\frac{\mathrm{d}}{\mathrm{d}m} \int_{r'}^{t'} U(s+m, s)\mathrm{d}s\bigg|_{m=0} = 0 \quad \text{for all } r' < t' \text{ in } (0,1).$$

Given $r < t$ in $(0,1)$, writing $\overline{\mathrm{D}}\phi_{r,t}(s') := \frac{\overline{\mathrm{d}}}{\mathrm{d}m}\phi_{r,t}(s'+m)\big|_{m=0}$ for the upper derivative, the outer Spence–Mirrlees condition yields for each $n \in (0,r)$ that

$$\begin{aligned} 0 &\leq \frac{\overline{\mathrm{d}}}{\mathrm{d}m} \int_{r-n}^{t-n} U(s+m, s+n)\mathrm{d}s\bigg|_{m=0} \\ &= \frac{\overline{\mathrm{d}}}{\mathrm{d}m} \int_{r}^{t} U(s+m-n, s)\mathrm{d}s\bigg|_{m=0} = \overline{\mathrm{D}}\phi_{r,t}(-n), \end{aligned}$$

which is to say that $\overline{\mathrm{D}}\phi_{r,t} \geq 0$ on $(-r,0)$. Since $\phi_{r,t}$ is continuous, it follows that $\phi_{r,t}$ is increasing on $[-r,0]$.[43] A similar argument shows that $\phi_{r,t}$ is decreasing on $[0,1-t]$.

It follows that for any $r < t$ in $[0,1]$ and $m \in [-r, 1-t]$,

$$\int_{r}^{t} [U(s,s) - U(s+m,s)]\mathrm{d}s = \phi_{r,t}(0) - \phi_{r,t}(m) \geq 0.$$

Thus for every $m \in [0,1]$, we have

$$U(s,s) - U(s+m,s) \geq 0 \quad \text{for a.e. } s \in [0,1] \cap [-m, 1-m].$$

Since $s \mapsto U(s,s) = V_{Y,P}(s)$ and $s \mapsto U(s+m,s)$ are continuous for any $m \in [0,1]$, it follows that for every $m \in [0,1]$,

$$U(s,s) - U(s+m,s) \geq 0 \quad \text{for } \textit{every } s \in [0,1] \cap [-m, 1-m],$$

which is to say that (Y,P) is incentive-compatible.

Step 2: Now drop the assumption that Y is continuous. By regularity of $\mathcal{Y}$ and the approximation lemma in appendix H.2, there exists a sequence $(Y_n)_{n \in \mathbf{N}}$ of continuous and increasing maps $[0,1] \to \mathcal{C}$ converging pointwise to Y, each of which satisfies $Y_n = Y$ on $\{0,1\}$. At each $n \in \mathbf{N}$, Step 1 yields a $P_n : [0,1] \to \mathbf{R}$ such that such that (Y_n, P_n) is incentive-compatible and satisfies the envelope formula with $V_{Y_n,P_n}(0) = 0$.

The sequence $(V_{Y_n,P_n})_{n \in \mathbf{N}}$ is Lipschitz equi-continuous[44] by the envelope formula and the boundedness of f_3. It is furthermore uniformly bounded, due to its Lipschitz equi-continuity and the fact that $V_{Y_n,P_n}(0) = 0$ for every $n \in \mathbf{N}$. Thus by the

[43] This is a standard result; see e.g. Bruckner (1994, §11.4, p. 128).

[44] That is, there is an $L \geq 0$ such that V_{Y_n,P_n} is L-Lipschitz for every $n \in \mathbf{N}$.

Arzelà–Ascoli theorem,[45] we may assume (passing to a subsequence if necessary) that $(V_{Y_n,P_n})_{n\in\mathbf{N}}$ converges pointwise. Then $(P_n)_{n\in\mathbf{N}}$ converges pointwise;[46] write $P : [0,1] \to \mathbf{R}$ for its limit.

By continuity of f, $U_n(r,t) \coloneqq f(Y_n(r), P_n(r), t)$ converges to $U(r,t) \coloneqq f(Y(r), P(r), t)$ for all $r, t \in [0,1]$. Each of the incentive-compatibility inequalities $U_n(t,t) \geq U_n(r,t)$ is preserved in the limit $n \to \infty$, ensuring that (Y, P) is incentive-compatible. ■

I Converse to the implementability theorem (§4.3, p. 13)

In this appendix, we provide a partial converse to the implementability theorem, and use it to prove Proposition 1 (p. 13). We shall use the partial converse again in appendix J below to prove Proposition 3 (p. 15).

Letting $\lesssim$ denote the partial order on $\mathcal{Y}$, we say that an allocation $Y : [0,1] \to \mathcal{Y}$ is *non-decreasing* iff there are no $t \leq t'$ in $[0,1]$ such that $Y(t') < Y(t)$. In other words, $Y(t)$ and $Y(t')$ could either be ranked as $Y(t) \lesssim Y(t')$, or they could be incomparable. Increasing maps are non-decreasing, but the converse is false except if $\mathcal{Y}$ is a chain.

Proposition 1′. If f is regular and satisfies the strict outer Spence–Mirrlees condition, then only non-decreasing allocations are implementable.

Proof of Proposition 1 (p. 13). By the implementability theorem, any increasing allocation is implementable. By Proposition 1′, any implementable allocation is non-decreasing, hence increasing since $\mathcal{Y}$ is a chain. ■

The proof of Proposition 1′ relies on two lemmata. The first is a 'non-decreasing' comparative statics result:[47]

Lemma 6. Let $\mathcal{X}$ and $\mathcal{T}$ be partially ordered sets, and let f be a function $\mathcal{X} \times \mathcal{T} \to \mathbf{R}$. Call a decision rule $X : \mathcal{T} \to \mathcal{X}$ *optimal* iff $f(X(t), t) \geq f(x, t)$ for all $x \in \mathcal{X}$ and $t \in \mathcal{T}$. If f has strictly single-crossing differences,[48] then every optimal decision rule is non-decreasing.

Proof. Write $\lesssim$ and $\preceq$, respectively, for the partial orders on $\mathcal{X}$ and on $\mathcal{T}$. Let $X : \mathcal{T} \to \mathcal{X}$ be optimal, and suppose toward a contradiction that there are $t \prec t'$

[45] E.g. Theorem 4.44 in Folland (1999).

[46] Clearly $f(Y_n(t), \inf_{m\geq n} P_m(t), t) = \sup_{m\geq n} f(Y_n(t), P_m(t), t) \leq \sup_{m\geq n} V_{Y_m,P_m}(t)$ for any $t \in [0,1]$, and thus $f(Y(t), p, t) \leq V(t)$, where $p \coloneqq \liminf_{n\to\infty} P_n(t)$ and $V(t) \coloneqq \lim_{n\to\infty} V_{Y_n,P_n}(t)$. Similarly $V(t) \leq f(Y(t), p', t)$, where $p' \coloneqq \limsup_{n\to\infty} P_n(t)$. Thus $f(Y(t), p, t) \leq f(Y(t), p', t)$, which rules out $p < p'$ since $f(Y(t), \cdot, t)$ is strictly decreasing.

[47] Such results are dimly known in the literature, but rarely seen in print. Exceptions include Quah and Strulovici (2007, Proposition 5) and Anderson and Smith (2021).

[48] A function $\phi : \mathcal{X} \times \mathcal{T} \to \mathbf{R}$ has *(strictly) single-crossing differences* iff $t \mapsto \phi(x', t) - \phi(x, t)$ is (strictly) single-crossing for any $x < x'$ in $\mathcal{X}$, where $<$ denotes the strict part of the partial order on $\mathcal{X}$. ('Single-crossing' was defined in footnote 17 on p. 12.)

in $\mathcal{T}$ such that $X(t') < X(t)$. Since $X(t)$ is optimal at parameter t, we have $f(X(t'),t) \leq f(X(t),t)$. Because $t \prec t'$ and $X(t') \prec X(t)$, it follows by strictly single-crossing differences that $f(X(t'),t') < f(X(t),t')$, a contradiction with the optimality of $X(t')$ at parameter t'. ■

Lemma 7. If f is regular and satisfies the (strict) outer Spence–Mirrlees condition, then for any price schedule $\pi : \mathcal{Y} \to \mathbf{R}$, the map $(y,t) \mapsto f(y,\pi(y),t)$ has (strictly) single-crossing differences.

Proof. Fix $y < y'$ in $\mathcal{Y}$, p, p' in $\mathbf{R}$ and $t < t'$ in $[0,1]$. Define a mechanism $(Y,P) : [0,1] \to \mathcal{Y} \times \mathbf{R}$ by $(Y(s),P(s)) \coloneqq (y,p)$ for $s \leq t$ and $(Y(s),P(s)) \coloneqq (y',p')$ for $s > t$, and fix $r, r' \in (0,1)$ with $r < t < r'$. Clearly for $n \in \{0, t'-t\}$,

$$\begin{aligned}
&\frac{\bar{\mathrm{d}}}{\bar{\mathrm{d}}m} \int_r^{r'} f\left(Y(s+m), P(s+m), s+n\right) \mathrm{d}s \bigg|_{m=0} \\
&\qquad\qquad = \frac{\mathrm{d}}{\mathrm{d}m} \left(\int_r^{t-m} f\left(y,p,s+n\right) \mathrm{d}s + \int_{t-m}^{r'} f\left(y',p',s+n\right) \mathrm{d}s \right) \bigg|_{m=0} \\
&\qquad\qquad = f\left(y',p',t+n\right) - f\left(y,p,t+n\right).
\end{aligned}$$

If f satisfies the outer Spence–Mirrlees condition, then the left-hand side is single-crossing in n, and thus $f\left(y',p',t\right) - f\left(y,p,t\right) \geq (>)\, 0$ implies $f\left(y',p',t'\right) - f\left(y,p,t'\right) \geq (>)\, 0$. Similarly for the strict case. ■

Proof of Proposition 1′. Let $Y : [0,1] \to \mathcal{Y}$ be implementable, so that (Y,P) is incentive-compatible for some payment schedule $P : [0,1] \to \mathbf{R}$. Define a price schedule $\pi : Y([0,1]) \to \mathbf{R}$ by $\pi \circ Y = P$; it is well-defined because by incentive-compatibility and strict monotonicity of $f(y,\cdot,t)$, $Y(r) = Y(r')$ implies $P(r) = P(r')$. Define a function $\phi : Y([0,1]) \times [0,1] \to \mathbf{R}$ by $\phi(y,t) \coloneqq f(y,\pi(y),t)$. Take any $t \in [0,1]$ and $y \in Y([0,1])$, and observe that there must be an $r \in [0,1]$ with $Y(r) = y$. Then since (Y,P) is incentive-compatible,

$$\begin{aligned}
\phi(Y(t),t) = f(Y(t),\pi(Y(t)),t) &= f(Y(t),P(t),t) \\
&\geq f(Y(r),P(r),t) = f(y,\pi(y),t) = \phi(y,t).
\end{aligned}$$

Since $y \in Y([0,1])$ and $t \in [0,1]$ were arbitrary, this shows that Y is an optimal decision rule for objective ϕ. Since ϕ has strictly single-crossing differences by Lemma 7, it follows by Lemma 6 that Y is non-decreasing. ■

J Proof of Proposition 3 (§4.4, p. 15)

Any increasing $Y : [0,1] \to \mathcal{Y}$ is implementable by the implementability theorem (§4.3, p. 13), and clearly sharing-proof. For the converse, let $Y : [0,1] \to \mathcal{Y}$ be implementable and sharing-proof, and fix $t < t'$; then either $Y(t) \lesssim Y(t')$ or $Y(t') < Y(t)$ since Y is sharing-proof, and it cannot be the latter because Y is non-decreasing by Proposition $1'$ in appendix I. ■

References

Anderson, A., & Smith, L. (2021). *The comparative statics of sorting* [working paper, 14 Jun 2021]. https://doi.org/10.2139/ssrn.3388017

Battaglini, M. (2005). Long-term contracting with Markovian consumers. *American Economic Review*, *95*(3), 637–658. https://doi.org/10.1257/0002828054201369

Benveniste, L. M., & Scheinkman, J. A. (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica*, *47*(3), 727–732. https://doi.org/10.2307/1910417

Blackwell, D. (1951). Comparison of experiments. In J. Neyman (Ed.), *Berkeley symposium on mathematical statistics and probability* (pp. 93–102). University of California Press.

Bruckner, A. (1994). *Differentiation of real functions* (2nd). American Mathematical Society.

Clausen, A., & Strub, C. (2020). Reverse calculus and nested optimization. *Journal of Economic Theory*, *187*. https://doi.org/10.1016/j.jet.2020.105019

Courty, P., & Li, H. (2000). Sequential screening. *Review of Economic Studies*, *67*(4), 697–717. https://doi.org/10.1111/1467-937X.00150

Danskin, J. M. (1966). The theory of max–min, with applications. *SIAM Journal on Applied Mathematics*, *14*(4), 641–664. https://doi.org/10.1137/0114053

Danskin, J. M. (1967). *The theory of max–min and its application to weapons allocation problems*. Springer.

Eső, P., & Szentes, B. (2007). Optimal information disclosure in auctions and the handicap auction. *Review of Economic Studies*, *74*(3), 705–731. https://doi.org/10.1111/j.1467-937x.2007.00442.x

Fitzpatrick, P. M., & Hunt, B. R. (2015). Absolute continuity of a function and uniform integrability of its divided differences. *American Mathematical Monthly*, *122*(4), 362–366. https://doi.org/10.4169/amer.math.monthly.122.04.362

Folland, G. B. (1999). *Real analysis: Modern techniques and their applications* (2nd). Wiley.

García, D. (2005). Monotonicity in direct revelation mechanisms. *Economics Letters*, *88*(1), 21–26. https://doi.org/10.1016/j.econlet.2004.12.022

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378. https://doi.org/10.1198/016214506000001437

Guesnerie, R., & Laffont, J.-J. (1984). A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics*, *25*(3), 329–369. https://doi.org/10.1016/0047-2727(84)90060-4

Hale, J. K. (1980). *Ordinary differential equations* (2nd). Krieger.

Hotelling, H. (1932). Edgeworth's taxation paradox and the nature of demand and supply functions. *Journal of Political Economy*, *40*(5), 577–616. https://doi.org/10.1086/254387

Kamenica, E., & Gentzkow, M. (2011). Bayesian persusasion. *American Economic Review*, *101*(6), 2590–2615. https://doi.org/10.1257/aer.101.6.2590

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

Matthews, S., & Moore, J. (1987). Monopoly provision of quality and warranties: An exploration in the theory of multidimensional screening. *Econometrica*, *55*(2), 441–467. https://doi.org/10.2307/1913245

Milgrom, P. (2004). *Putting auction theory to work*. Cambridge University Press. https://doi.org/10.1017/CBO9780511813825

Milgrom, P., & Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, *70*(2), 583–601. https://doi.org/10.1111/1468-0262.00296

Milgrom, P., & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, *62*(1), 157–180. https://doi.org/10.2307/2951479

Mirrlees, J. A. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics*, *6*(4), 327–358. https://doi.org/10.1016/0047-2727(76)90047-5

Morand, O., Reffett, K., & Tarafdar, S. (2015). A nonsmooth approach to envelope theorems. *Journal of Mathematical Economics*, *61*, 157–165. https://doi.org/10.1016/j.jmateco.2015.09.001

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, *6*(1), 58–73. https://doi.org/10.1287/moor.6.1.58

Nöldeke, G., & Samuelson, L. (2018). The implementation duality. *Econometrica*, *86*(4), 1283–1324. https://doi.org/10.3982/ECTA13307

Oyama, D., & Takenawa, T. (2018). On the (non-)differentiability of the optimal value function when the optimal solution is unique. *Journal of Mathematical Economics*, *76*, 21–32. https://doi.org/10.1016/j.jmateco.2018.02.004

Pavan, A., Segal, I., & Toikka, J. (2014). Dynamic mechanism design: A Myersonian approach. *Econometrica*, *82*(2), 601–653. https://doi.org/10.3982/ECTA10269

Quah, J. K.-H., & Strulovici, B. (2007). *Comparative statics with the interval dominance order: Some extensions* [working paper].

Roy, R. (1947). La distribution du revenu entre les divers biens. *Econometrica*, *15*(3), 205–225. https://doi.org/10.2307/1905479

Royden, H. L., & Fitzpatrick, P. M. (2010). *Real analysis* (4th). Prentice Hall.

Samuelson, P. A. (1947). *Foundations of economic analysis*. Harvard University Press.

Shaked, M., & Shanthikumar, J. G. (2007). *Stochastic orders*. Springer.

Shephard, R. W. (1953). *Cost and production functions*. Princeton University Press.

Silberberg, E. (1974). A revision of comparative statics methodology in economics, or, how to do comparative statics on the back of an envelope. *Journal of Economic Theory*, *7*(2), 159–172. https://doi.org/10.1016/0022-0531(74)90104-5

Spence, M. (1974). Competitive and optimal responses to signals: An analysis of efficiency and distribution. *Journal of Economic Theory*, *7*(3), 296–332. https://doi.org/10.1016/0022-0531(74)90098-2

Teschl, G. (2012). *Ordinary differential equations and dynamical systems*. American Mathematical Society.