THE UNIVERSITY OF
MELBOURNE

Department of Economics

Working Paper Series

Duality in Contracting

Peter Bardsley

February 2012

Research Paper Number 1141

ISSN: 0819-2642
ISBN: 978 0 7340 4491 4

Department of Economics
The University of Melbourne
Parkville VIC 3010
www.economics.unimelb.edu.au

Duality in Contracting

Peter Bardsley
The University of Melbourne

February 2012

Abstract

In a linear contracting environment the Fenchel transform provides a complete duality between the contract and the information rent. Through an appropriate generalised convexity this can be extended to provide a complete duality in the supermodular quasilinear contracting environment that covers the majority of applications. Using this framework, we provide a complete characterization of the allocation correspondences that can be implemented by a principal in this environment. We also address the question of when an allocation can be implemented by a menu of simple contracts. Along the way, a supermodular envelope theorem is proved, somewhat different in nature to the Milgrom Segal result.

Key Words: mechanism design, contract theory, duality, Fenchel transform, abstract convexity.

JEL Codes: D82, D86

1 Introduction

In the linear agency problem, that is to say when the valuation is linear in type, the agent is confronted with a decision problem

$$\max_{x \in X} \theta x - \tau(x).$$

Here θ is the agent's type, $x \in X$ is the allocation chosen by the agent, and $\tau(x)$ is the contract or tariff designed by the principal. This problem is identical in structure to the standard producer problem in price theory

$$\max_{x \in X} px - c(x)$$

where p is the price, $x \in X$ is the output quantity chosen by the agent, and $c(x)$ is the cost function.

The duality framework of price theory embeds the producer's problem as one of a pair of dual problems. The primal problem, described by the cost function, lives on the commodity space and the dual problem, described by the profit function, lives on the price space. These two structures are inter-linked by the envelope theorem and its variants. This framework clarifies the producer problem and underlies many useful technical results, in particular the regularity and convexity of the profit function. The duality transform of price theory is a particular example of the Fenchel transform, which is studied systematically in convex analysis (Rockafellar 1970).

Since the linear agency problem and the producer problem are isomorphic, there is a completely parallel duality framework for the linear agency problem, with the type θ playing the role of the price p. The primal problem is described by the contract function, living on the allocation space, and the dual problem is described by the information rent, living on the type space. While this observation is completely elementary, it is already useful and not always fully evident in standard treatments of the theory. It is exploited, for example, by Krishna and Maenner (2001) in auction theory.

In this paper we consider the more general quasi-linear adverse selection contracting problem where an agent of type $\theta \in \Theta$ chooses $\xi(\theta) \subset X$ to solve

$$\rho(\theta) = \sup_{x \in X} v(\theta, x) - \tau(x) \tag{1a}$$

$$= \sup_{(x,t) \in \Gamma} v(\theta, x) - t \tag{1b}$$

$$\xi(\theta) = \operatorname*{argmax}_{x \in X} v(\theta, x) - \tau(x) \tag{1c}$$

$$= \operatorname*{argmax}_{(x,t) \in \Gamma} v(\theta, x) - t. \tag{1d}$$

Here X and Θ are intervals in $\mathbb{R}$, $\Gamma = \operatorname{epi} \tau = \{(x,t) : t \geq \tau(x)\}$ is the epigraph of the contract function $\tau : X \to \bar{\mathbb{R}}$, and $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ is the extended real line, with ∞ being an element greater than any real number. There is no assumption that either X or Θ is compact. As is standard in convex analysis, allowing extended real values is a convenient way of restricting the domain of τ. Points where $\tau(x) = \infty$ are infeasible, since the transfer $\tau(x)$ is infinite. It will be assumed that τ is proper, that is, $\operatorname{dom}(\tau) = \{x : \tau(x) < \infty\} \neq \emptyset$ in order to exclude uninteresting vacuous cases. Assumptions on the agent's valuation function $v(x, \theta)$ will be discussed below.

A standard interpretation of this problem is that $x \in \xi(\theta)$ is a quantity sold by a principal to an agent of type θ in return for a payment $\tau(x)$. Equivalently, Γ is a menu of contracts offered by the principal. Under this interpretation the value function $\rho(\theta)$ is the information rent accruing to the agent of type θ. There are of course many other interpretations of this canonical model.

The main contribution of this paper is to show that the linear duality framework can be extended, virtually in its totality, to the quasilinear framework that is standard in most applications of contract theory.[1] The extension requires a non-trivial but quite intuitive generalisation of the concept of convexity, which is of interest in itself because it links convexity (a geometrical property) with supermodularity (an incentive property).

The key idea of convex duality is that a convex set can be approximated by the affine half spaces that support it. Broadly speaking, a set is amenable to such approximation if it has no concave sections — hollows or dents that cannot be penetrated by a supporting affine hyperplane. Nonconvex duality extends this idea by using a family of primitive objects, for example conical or parabolic sets, that can penetrate hollows and approximate a larger family of objects. The main innovation in this paper is to show that a satisfactory duality can be constructed by building these primitive approximating objects not in an ad hoc way but from the problem at hand: from the agent's indifference curves in contract space. It will transpire that a satisfactory nonconvex duality can be defined if and only if the agent's valuation $v(x, \theta)$ is supermodular (after perhaps relabeling some types). This observation provides an interesting explanation of why it is that supermodularity or single crossing assumptions so often lead to tractable models.

Just as in the classical case, generalised convexity leads to strong regularity properties for envelope functions. We prove a non-convex separation theorem, and display a full duality between the contract and the information rent. In particular, not only the information rent but also the contract function can be represented as an integral over marginal valuations. Along the way we prove an envelope theorem that is related to but different from the envelope theorem of Milgrom and Segal (2002). The envelope theorem proved here is particularly well adapted to the quasilinear environment, where it is easier to apply and produces a stronger result. Using these tools we generalize a classical result of Rochet (1987) to provide a complete characterization of the allocation correspondences that are implementable by the principal in

[1] We assume throughout that the type space is 1 dimensional.

a quasilinear contracting environment.

The focus in this paper is on the agent's problem. Relatively little will be said about the principal's problem or about the structure of optimal contracts. The approach is relatively direct, in so far as there will be no use made of control theory. In particular there is no need for any a priori assumption that the agent's choice $x(\theta)$ is an absolutely continuous function or for delicate arguments in control theory, as is required in most standard control theoretic approaches to the agency problem. There will be no discussion of participation constraints or individual rationality, as the agent's participation decision can be separated from the agent's action decision conditional on participation, which is what we study here. The participation decision is straight forward, as it is monotonic in type, and the ideas presented here throw no new light on this.

The results in this paper are related to those of Rochet (1987), who characterizes the allocation functions that are implementable in a quasilinear contracting environment (we characterise the class of implementable correspondences), to Krishna and Maenner (2001), Border (1991), Border (2007) and Vohra (2011) who exploit linear duality and convex geometry in auction theory, and to Milgrom and Segal (2002). It is also related to the literature on nonsmooth analysis, generalised convexity and the generalised Fenchel transform, which is comprehensively addressed in Rockafellar and Wets (1998).

The structure of the paper is as follows. Section 2 sets out the structure of the duality that arises in the linear agency case (that is, the agent's valuation is linear; the contract is of course not restricted to be linear). The techniques used here are those of classical convex analysis. The results in this Section generalize, essentially in their totality, to the quasilinear case with unchanged proofs once the machinery is developed later in the paper. It is useful, however, to develop them first in the linear environment since this makes the structure of the results much clearer. In Section 3 the necessary machinery is developed to support an appropriate non-convex duality and a generalised Fenchel transform. Lemma 1 in this section, setting out the link between supermodularity and generalised convexity, is perhaps the conceptual heart of the paper. The technical heart is Theorem 2, which generalizes the classical result that a function is an affine envelope if and only if it is proper, convex, and lower semicontinuous. Important regularity results, including an envelope theorem, are also established in this Section. Once this machinery is in place, the duality results for the quasilinear environment follow immediately. These are set out in Section 4, in which we also address implementation by menus of simple contracts. Section 5 is the conclusion. Several of the longer proofs are relegated to an appendix.

2 Linear Agency

In this section we study Problem (1) under the assumption that the agent's valuation $v(x, \theta) = \theta x$ is linear in both the quantity x and the agent's type θ. When we refer to this as a *linear agency problem* we mean that the agent's valuation θx is linear; the contract $\tau(x)$ will of course not in general be linear. It is useful to explore the linear case in some detail as a first step because the structure of the results can be seen with a minimum of technical machinery. This structure, and most of the proofs, will carry across virtually unchanged to the general case, once some machinery is developed.

There are two ways to approach the concept of convexity: internal and external. The internal route characterizes convexity of a set by the property that if the set contains two points then it must contain the interval between them. The external route approaches convexity by requiring that a set be the intersection of the affine half spaces containing it. Of course this leads to a slightly different concept: in finite dimensions a set is the intersection of affine half spaces if and only if it is both closed and convex. In particular, it is a fundamental result that if a proper function is bounded below[2] then

[2] It is sufficient that it be bounded below by some affine function.

it is the envelope of its affine supports if and only if it is convex and lower semicontinuous. We will call such functions *envelope functions.* The interplay between these two convexity concepts, internal and external, generates many of the important results in convexity and will be fundamental to this paper as well.

It is useful to recall some basic definitions and results from convex analysis, as a basis for generalizations that will be presented later (for details, see Rockafellar (1970), Bonnans and Shapiro (2000) or Rockafellar and Wets (1998)). A proper function $\tau(x)$ is convex if its epigraph $\operatorname{epi}\tau$ is a convex set. It is lower semicontinuous if $\operatorname{epi}\tau$ is a closed set. An affine function $\phi(z) = \alpha z + \beta$ is a minorant of τ if its graph lies below $\operatorname{epi}\tau$; that is, $\phi(z) \leq \tau(z)$ for all z. An affine support at x is an affine minorant ϕ whose graph touches $\operatorname{epi}\tau$ at x; that is, $\phi(x) = \tau(x)$. The affine envelope $\tilde{\tau}(x)$ of $\tau(x)$ is the supremum of the affine minorants of $\tau(x)$. It is an envelope function less than or equal to $\tau(x)$, and it is the greatest such function.

Given a proper convex function $\tau(x)$, its Fenchel transform or conjugate is

$$\rho(\theta) = \tau^*(\theta) = \sup_{x\in X} \theta x - \tau(x) \tag{2a}$$

$$\xi(\theta) = \operatorname*{argmax}_{x\in X} \theta x - \tau(x) \tag{2b}$$

$$\eta(x) = \xi^{-1}(x) = \{\theta \in \Theta : (\theta, x) \in \xi\}. \tag{2c}$$

The interpretation of these equations in a contracting environment is that $\tau(x)$ is the contract function, $\rho(\theta) = \tau^*(\theta)$ is the information rent function (we will use these notations interchangeably, depending on whether we have in mind a contracting framework or a more abstract framework), $\xi(\theta)$ is the set of optimal allocations from which an agent of type θ can choose, and $\eta(x)$ is the set of types who can optimally choose x. We will let $x(\theta) \in \xi(\theta)$ and $\theta(x) \in \eta(x)$ be selections from the respective correspondences.

The conjugate τ^* has attractive properties. For a start, it is by construction the envelope of its affine minorants, so it is an envelope function. It is convex and lower semicontinuous. Furthermore it is absolutely continuous and locally Lipshitz continuous on the interior of its domain, differentiable almost everywhere, and it is the integral of its derivative. The allocation correspondence ξ is monotone: if $\theta' \geq \theta$ and $x' \in \xi(\theta')$, $x \in \xi(\theta)$ then $x' \geq x$.

The most important case is when the allocation ξ is not just monotone but *maximal monotone* (that is, ξ cannot be expanded as a correspondence without violating monotonicity). This occurs precisely when the contract function $\tau(x)$ is an envelope function (the information rent $\rho(\theta)$ is always, by construction, an envelope function). That is to say, in the linear environment these maximal monotone allocation correspondences are precisely those that can be implemented by envelope contracts (Rockafellar and Wets (1998, 12.26)). The study of the case where ξ is monotone but not maximal can be reduced to the maximal monotone case, as will be discussed below. Maximal monotone correspondences on $\mathbb{R}$ have a simple characterization that we will use repeatedly. Let $x(\theta) \in \xi(\theta)$ be a selection from the correspondence. Then $x(\theta)$ is a monotonic function so it is continuous except at a countable number of points where it has a jump discontinuity. A maximal monotone correspondence ξ can be reconstructed from the graph of such a monotone function $x(\theta)$ by "filling in the jumps". The correspondence ξ, considered as a set valued function, is given by the formula $\xi(\theta) = [x_-(\theta), x_+(\theta)]$, where $x_-(\theta)$ and $x_+(\theta)$ are the limits of $x(\theta)$ from the left and the right respectively.

Just as in classical price theory, when the allocation ξ is maximal monotone, Fenchel conjugation leads to a complete duality between the contract τ and the information rent ρ. In particular we have the following properties.[3] Both

[3]The only standard result that is missing here, and which does not generalize to the quasilinear case, is the identification of $\xi(\theta)$ and $\eta(x)$ with the subgradients $\partial\rho(\theta)$ and $\partial\tau(x)$. It seems likely that this could be done using one of the many nonsmooth generalizations of the classical subgradient of convex analysis.

τ and ρ are envelope functions; they are convex and lower semicontinuous on X and Θ respectively; they are continuous on their domains,[4] locally Lipshitz continuous and absolutely continuous on the interior of their domains. τ and ρ are supported by their affine minorants on the interior of their domains (that is, they are not just the supremum of their affine minorants — the maximum is actually attained). τ and ρ are Fenchel conjugates: $\rho = \tau^*$ and $\tau = \rho^*$. In particular, $\tau = \tau^{**}$ and $\rho = \rho^{**}$. The correspondences ξ and η are maximal monotone. The Fenchel inequality: $\tau(x) + \rho(\theta) \geq \theta x$ holds, with equality iff $\theta \in \eta(x)$ iff $x \in \xi(\theta)$. $\tau(x)$ and $\rho(\theta)$ are differentiable almost everywhere, and $\tau'(x) \in \eta(x) = \{\tau'(x)\}$, $\rho'(\theta) \in \xi(\theta) = \{\rho'(\theta)\}$ at points of differentiability. In this case both the contract $\tau(x)$ and the rent $\rho(\theta)$ have integral representations

$$\rho(\theta) = \rho(\theta_0) + \int_{\theta_0}^{\theta} x(\theta)\, d\theta \tag{3a}$$

$$\tau(x) = \tau(x_0) + \int_{x_0}^{x} \theta(x)\, dx \tag{3b}$$

where $x(\theta) \in \xi(\theta)$ and $\theta(x) \in \eta(x)$ are selections from the respective correspondences. The integral envelope representation of the information rent $\rho(\theta)$ is standard. The entirely symmetric relationship for the contract function $\tau(x)$ is striking, but often unnoticed.

If ξ is not maximal monotone then $\tau(x)$ is not an envelope function. It is natural in this case to consider the affine envelope $\tilde{\tau}(x) = \tau^{**}(x)$ generated by $\tau(x)$, and the maximal monotone allocation $\tilde{\xi}(\theta)$ implemented by $\tilde{\tau}(x)$. The affine envelope $\tilde{\tau}(x)$ is the supremum of the affine minorants of $\tau(x)$, it is the greatest envelope function that minorises $\tau(x)$, and it gives rise to the same information rent: $\rho(\theta) = \tau^*(\theta) = \tilde{\tau}^*(\theta)$. Hence $\tilde{\tau}(x)$ and $\rho(\theta)$ are Fenchel duals, tied together by duality relationships as described above. The

[4] That is, on the sets $\{\{x \in X : \tau(x) < \infty\}\}$ and $\{\{\theta \in \Theta : \rho(\theta) < \infty\}\}$. We use here the fact that X and Θ are 1 dimensional. In higher dimensions we get continuity only on the interior of the domain (see Rockafellar (1970, Chapter 10)).

connection between ξ, the allocation supported by τ, and $\tilde{\xi}$, the allocation supported by $\tilde{\tau}$, will be addressed below. We note immediately, however, that $\xi \subset \tilde{\xi}$ and $\eta \subset \tilde{\eta}$, that $\eta(x) = \emptyset$ if $\tau(x) > \tilde{\tau}(x)$, and that $\eta(x) = \tilde{\eta}(x)$ if $\tau(x) = \tilde{\tau}(x)$. This is so since τ and $\tilde{\tau}$ have the same affine minorants, and $\tilde{\tau} \leq \tau$, and if an affine of slope θ supports τ at x then it supports $\tilde{\tau}$ at x.

It thus follows that if $\tau(x) > \tilde{\tau}(x)$ then the exact value of $\tau(x)$ is immaterial. In economic terms this corresponds to the fact that in non-convex parts of the contract, which are inaccessible to supporting affine functions, the precise value of $\tau(x)$ is not important. It is natural to impose a normalization convention, which will not change the decision of any agent, that $\tau(x) = \infty$ if $\tau(x) > \tilde{\tau}(x)$. We will impose this convention from now on.

2.1 Implementable allocations

We have seen above that if ξ is the allocation implemented by a contract τ, then ξ is a sub-correspondence of the maximal monotone correspondence $\tilde{\xi}$ implemented by the affine envelope $\tilde{\tau}$. As a first application of this idea, we address the question of how to characterize the correspondences ξ that can be implemented in this linear environment. This implementation question has been addressed for functions, not correspondences, by Rochet (1987). For maximal monotone correspondences in a linear environment it is addressed by a classical result of Rockafellar (see Rockafellar and Wets (1998, 12.17)). We address the issue here for monotone (but not necessarily maximal monotone) correspondences in the linear environment. The result, and the proof, will generalise without change to the supermodular quasilinear environment once approriate machinery is developed later in the paper.

The following notation will be useful in stating and proving the result. It may also be useful to refer to Figure 1, and to the example below, which illustrate the construction. Let $\xi \subset \Theta \times X$ be a monotone allocation correspondence. As above we write $\eta = \xi^{-1}$ for the inverse correspondence, so $\xi(\theta)$ is the set of x that θ can choose under this allocation, and $\eta(x)$ is the set of types θ that can choose x. Note that $\operatorname{im}\xi = \operatorname{dom}\eta = \{x : \exists\theta : (\theta, x) \in \xi\}$ is the projection of the correspondence ξ onto X, and $\operatorname{dom}\xi = \operatorname{im}\eta = \{\theta : \exists x : (\theta, x) \in \xi\}$ is the projection onto Θ. Since ξ is monotone it is totally ordered under the standard partial order on $\Theta \times X$ (Rockafellar and Wets 1998).[5] A gap in ξ is a partition of $\xi = \xi_- \cup \xi_+$ into disjoint pieces with $\xi_- < \xi_+$ such that there exists $(\theta, x) \notin \xi$ with $\xi_- < (\theta, x) < \xi_+$. So ξ could be enlarged by adding (θ, x) without destroying monotonicity. Given such a gap, we write $(\theta_-, x_-) = \sup\xi_-$, and $(\theta_+, x_+) = \inf\xi_+$ for the lower and upper bounds to this gap. We allow, in this notation, that ξ_+ could be empty — we refer to this as a gap at the top, and we make a similar definition at the bottom. This notation requires a little care at the top and at the bottom of ξ. If there is a gap at the top then the supremum of ξ is (θ_-, x_-); this is the infimum of the gap at the top. Similarly, if there is a gap at the bottom then the infimum of ξ is (θ_+, x_+).

Theorem 1 *Let $\xi \subset \Theta \times X$ be correspondence. Then ξ is implementable in the linear contracting environment if and only if*

1. *ξ is monotone;*
2. *$\xi(\theta)$ is convex for all $\theta \in \Theta$;*
3. *$\eta(x)$ is closed and convex for all $x \in X$; and, at any gap $\xi_- < \xi_+$,*
4. *either $\eta(x_-) = \emptyset$, or $\eta(x_+) = \emptyset$ (both may be empty);*
5. *if $\eta(x_+) \neq \emptyset$ then $\theta_+ \in \eta(x_+)$; if the gap is at the bottom then $(-\infty, \theta_+] \subset \eta(x_+)$;*
6. *if $\eta(x_-) \neq \emptyset$ then $\theta_- \in \eta(x_-)$; if the gap is at the top then $[\theta_-,\infty) \subset \eta(x_+)$.*

[5] That is, $(\theta, x) \leq (\theta', x')$ if and only if $\theta \leq \theta'$ and $x \leq x'$.

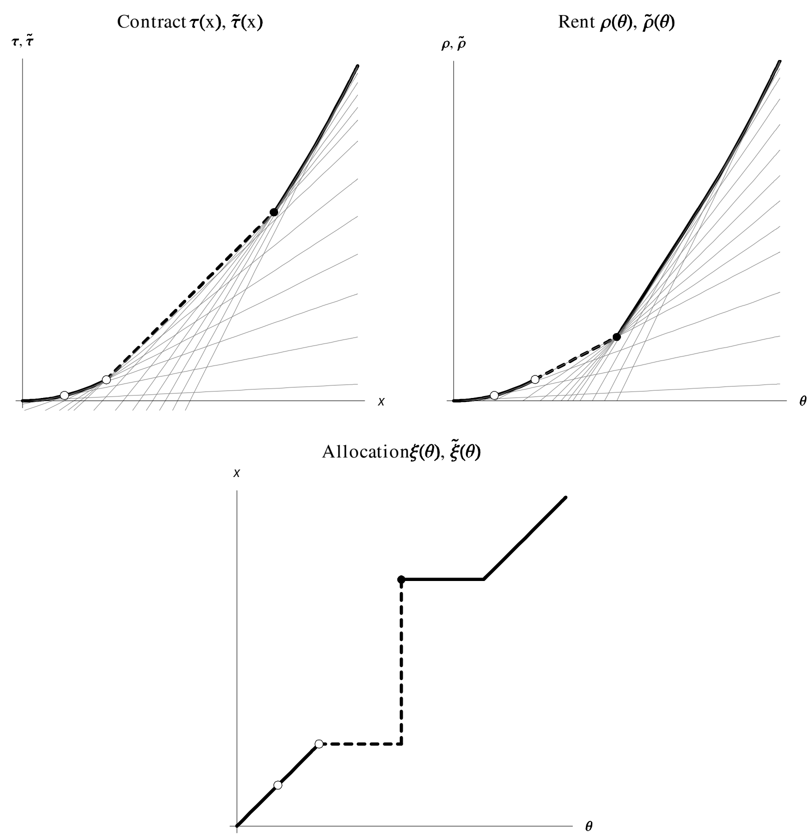

Figure 1: Implementation of the contract ξ (shown bold). $\tilde{\xi}$ (shown dashed) is a maximal monotone extension. $\tilde{\tau}$ is the envelope contract that supports $\tilde{\xi}$. The contract that supports ξ is the restriction τ of $\tilde{\tau}$. $\tilde{\rho}$ and ρ are information rents.

The sets ξ_-, ξ_+ need not be closed; conditions (5) and (6) can be regarded as a weak substitute for closure. Condition (5) , for example, says that if some type chooses x_+ then ξ_+ contains its limit point (θ_+, x_+) . For the proof, see the Appendix. The idea of the proof is as follows. Starting with the monotone allocation ξ, expand it to a maximal monotone correspondence $\bar{\xi}$. Using the integral representation discussed above (the integral form of the classical Fenchel envelope duality), construct a contract $\bar{\tau}$ that implements $\bar{\xi}$. Provided that the expansion $\bar{\xi}$ is constructed appropriately one can then check that the restriction of $\bar{\tau}$ to $\operatorname{dom}\xi$ implements ξ.

In the unidimensional case, under differentiability and single crossing assumptions on the valuation function $v(x,\theta)$ that certainly hold in the linear environment considered in this section, Rochet (1987) verifies that his necessary and sufficient condition amounts to monotonicity of the allocation function[6]. So, in the case that the allocation correspondence $\xi(\theta)$ is in fact a globally defined single valued function $x(\theta)$, our result coincides with that of Rochet (and will continue to do so in the quasilinear supermodular case). It also agrees with a classical result of Rockafellar in the linear environment if the allocation $\xi(\theta)$ is maximal monotone.

If $\xi(\theta)$ is maximal monotone then the supporting contract $\tau(x)$ and the information rent $\rho(\theta)$ are, by Fenchel duality, unique up to an additive constant (this is a form of revenue equivalence). If $\xi(\theta)$ is not maximal monotone then, as is clear from the proof, there is in general some latitude in constructing a supporting contract $\tau(x)$. There may be more than one way to enlarge ξ to a maximal monotone correspondence $\tilde{\xi}$. In particular, we may well choose a more general interpolation of ξ between (θ_-,x_-) and (θ_+,x_+) than the step function used in the construction above. In this case the uniqueness that underlies revenue equivalence may break down. As a result of this observation we have the following result.

Corollary 1 (revenue equivalence) *Let ξ be an allocation that is implementable in the linear contracting environment. Then the supporting contract τ and the information rent ρ are uniquely determined, up to an additive constant, if and only if there are no gaps $\xi_+ \geq \xi_-$, with $\xi_+,\xi_- \neq \emptyset$, such that $x_+ > x_-$ and $\theta_+ > \theta_-$.*

To illustrate this result consider the following example, which is illustrated in Figure 1. The allocation is defined by $\xi(\theta) = \{\theta\}$ if $0 \leq \theta < \frac{1}{2}$, $\frac{1}{2} < \theta < 1$, or $\theta > 3$, $\xi(\theta) = \{3\}$ if $3 \leq \theta < 3$, and $\xi(\theta) = \{\}$ otherwise. This is

[6] This observation dates back of course to the earliest literature in mechanism design.

shown in bold in Figure 1. A maximal monotone extension $\tilde{\xi}$ is shown dashed. There is considerable freedom of choice in constructing ξ; we use here a combination of horizontal and vertical steps as in the proof of Theorem 1. Here the vertical step is as far as possible to the right in the gap: amongst the contracts that support ξ this will generate the one that maximizes the agent's rent ρ. The contract that supports $\tilde{\xi}$ is $\tilde{\tau}(x) = \frac{x^2}{2}$ if $x \leq 1$, or $x \geq 3$, and $\tilde{\tau}(x) = 2x - \frac{3}{2}$ if $1 \leq x \leq 3$, yielding information rent $\tilde{\rho}(\theta) = \frac{\theta^2}{2}$ if $0 \leq \theta \leq 1$, or $\theta \geq 3$, $\tilde{\rho}(\theta) = \theta - \frac{1}{2}$ if $1 \leq \theta < 2$, and $\tilde{\rho}(\theta) = 3\theta - \frac{9}{2}$ if $2 < \theta \leq 3$. The contract τ that supports ξ is the restriction of $\tilde{\tau}$ to $\operatorname{im} \xi = \left\{x : 0 \leq x < \frac{1}{2},\ \frac{1}{2} < x < 1,\ \text{or } x \geq 3\right\}$. The information rent induced by this contract is $\rho(\theta) = \frac{\theta^2}{2}$ if $0 \leq \theta < \frac{1}{2}$, $\frac{1}{2} < \theta < 1$, or $\theta \geq 3$.

2.2 Will the principal choose an envelope contract?

Is it reasonable to focus on envelope contracts? Is there any loss of generality in such an assumption? In price theory, it is commonly argued that the predictive content of the theory is unchanged if a cost function $c(x)$ is replaced with its lower semicontinuous convexification $\tilde{c}(x)$. That is to say, there is no significant loss of generality in assuming that $c(x)$ is an envelope function. The argument is that points on the cost function where $c(x) \neq \tilde{c}(x)$ are economically irrelevant, since they are inaccessible to any hyperplane and they will never be chosen at any price. For an exposition of this point of view, set out with respect to consumer theory rather than producer theory, see Jehle and Reny (2000, Chapter 2.1.2). This argument suggests that without significant loss of generality we may assume that the principal will offer an envelope contract $\tau(x)$. And it is certainly the case that optimal contracts constructed using the control theoretic techniques typical in applied mechanism design will be envelope contracts. The following rather preliminary result reassures us that in a large class of examples the principal will indeed choose an envelope contract. A more thorough investigation of the nature of optimal contracts requires ideas beyond those in this paper.

Proposition 1 *Assume that the Principal chooses a contract $\tau(x)$ to maximize the expected value of a continuous function $\psi(x,t)$ of the allocation and the payment, and that the distribution of types is non-atomic and has full support.*[7] *Then without loss of generality she will choose an envelope contract.*

Proof. It will be sufficient to show that she would be willing to choose the envelope contract $\tilde{\tau}$. That is to say, should would not gain by deleting points from the domain of $\tilde{\tau}$, as would be required if she offers $\tau \subset \tilde{\tau}$.

Let $\phi(\theta,x) = \psi(x,\tilde{\tau}(x))$. Note that this is continuous on $\tilde{\xi}$,the allocation supported by $\tilde{\tau}$, since $\tilde{\tau}$ is continuous on its domain, and without loss of generality $\phi(\theta,x)$ can be assumed to be non-negative on ξ. For let $Z = \{x \in \operatorname{dom}\tau : \phi(\theta,x) = \psi(x,\tilde{\tau}(x)) = \psi(x,\tau(x)) < 0\}$. If the principal deletes Z from $\operatorname{dom}\tau$ then an agent who is not choosing an allocation in Z will not change their behavior. An agent choosing an allocation in Z will either switch to an allocation not in Z, or they will no longer have an optimal choice in X and they will not participate[8]; in either case the payoff to the principal does not decrease. Thus we can assume that $Z = \emptyset$.

Consider the allocation $\tilde{\xi}$, supported by the contract $\tilde{\tau}$. Consider $(\theta,x) \in \tilde{\xi}$ with $x \notin Z$. Then (θ,x) lies in a gap in ξ, as discussed in the previous section. It is clear that $\tilde{\tau}$ is affine in this gap, and $\tilde{\xi}$ is a step correspondence, either θ is a mixing type with $\xi(\theta)$ being a non-degenerate mixing interval, or x is a limit point of K. Consider the latter case, and assume that $x = x_-$ is at the bottom of the gap (the case at the top of a gap is similar). Then by Theorem 1 $(\theta_-,x_-) \in \xi$, and $\phi(\theta_-,x_-) \geq 0$ by continuity of ϕ on ξ. But $\phi(\theta,x) = \phi(\theta,x_-) = \phi(\theta_-,x_-)$ since ϕ does not depend on θ. Thus $\phi(\theta,x) \geq 0$, and the principal cannot gain by deleting such points.

It remains to consider points x that lie in a non-degenerate mixing interval of $\tilde{\xi}$. If we delete such a point then we cannot argue that this would not change

[7] That is, no non-empty open interval in X has measure zero.

[8] That is to say, $\xi(\theta) = \emptyset$. This can occur if $\operatorname{dom}\tau$ is not closed, and the agent's optimal choice is at a limit point not in $\operatorname{dom}\tau$.

the behavior of the agent or the payoff to the principal from contracting with such an agent. But the set of mixing types is just the set of jump points of the monotone correspondence $\tilde{\xi}$. This set is countable, hence of measure zero since the type distribution is non-atomic. So deleting these points will not improve the payoff to the principal. ■

The assumption made here about the principal's objective is that she cares about the transacted quantity and the price, but not otherwise about the identity or type of the agent. This assumption might fail if, for example, the principal were altruistic or if higher types were more expensive to serve. In this case it might be optimal to create gaps in the allocation and exclude some types. But one must remark that any such non-envelope contract would have the somewhat implausible feature that a positive measure of agents fail to accept the contract not because they have a better outside option but only because there are contract points arbitrarily close to their optimal point, but no actual optimal response.

3 Generalised Convexity

The basic idea that drives convex duality is to approximate a set (or a function) by a family of primitive objects: affine hyperplanes. Broadly speaking, a set is amenable to such approximation if it has no concave sections — hollows or dents that cannot be penetrated by a supporting affine hyperplane. Nonconvex duality extends this idea by using a family of primitive objects, for example conical or parabolic sets, that can penetrate hollows and approximate a larger family of objects. See for example Rockafellar and Wets (1998, Chapter 11.K). This is the approach that will be adopted here in constructing a duality for contracting problems. The primitive approximating objects will not, however be arbitrary. They will be adapted to the problem at hand, and will be constructed from the payoff functions for the agents in contract space. Rather than begin with sets, it is more convenient to start with an abstract convexity concept for functions. Convexity for sets will then follow by requiring that their characteristic function be convex.

We now leave the linear case and return to the general quasilinear framework of Equations ($1a$) to ($1d$). Let Φ be the set of functions $f : \mathbb{R} \to \mathbb{R}$ of the form $f(z) = v(z, \theta) - \alpha$, where $\theta \in \Theta$, and $\alpha \in \mathbb{R}$. Following Rockafellar and Wets (1998) (see also Pallaschke and Rolewicz (1997) and Rubinov (2000)), regard Φ as a set of elementary functions which will be used to define an abstract convexity class. If Φ is the set of affine functions then this will lead to the class of convex functions of classical convex analysis (Rockafellar 1970), but we will work with an abstract convexity class that is more closely adapted to Problem 1. When we discuss duality below, it is useful to consider not only the elementary functions on X, but also the functions (which will also be called elementary) $g(z) = v(x, z) - \alpha$ on Θ. By a slight abuse of notation, Φ will be used to denote either class of elementary functions.

The existing literature moves directly from Φ to the class of *externally* Φ-convex functions, the envelopes of elementary functions in Φ, and does not address the issue of an appropriate definition of *internal* convexity or the relationship between the two ideas (Pallaschke and Rolewicz (1997) and Rubinov (2000)). In our context we have sufficient structure to address this issue properly, and there are significant payoffs from doing so. The basic construct is to define a Φ-interval in $\mathbb{R} \times \mathbb{R}$. This is a set of the form $\{(x, v(x, \theta) - \alpha) : x \in [x_0, x_1]\}$, where $\theta, \alpha \in \mathbb{R}$ and $[x_0, x_1]$ is an interval in $\mathbb{R}$. A Φ-interval in $\mathbb{R} \times \mathbb{R}$ is thus the graph of an elementary Φ-function $v(x, \theta) - \alpha$ restricted to an interval $[x_0, x_1]$. If we note that the equation for the indifference curve for an agent of type θ is $t = v(x, \theta) - \alpha$, then it is apparent that a Φ-interval is just a connected segment of an agent's indifference curve in contract space. A Φ-interval joining two points plays the same role in Φ-convexity as rectilinear intervals play in classical convexity. A set $S \subset \mathbb{R} \times \mathbb{R}$ is *internally* Φ*-convex* if, for any $(x, t), (x', t') \in S$, and any Φ-interval I with endpoints (x, t) and (x', t'), $I \subset S$. It may be useful to note that in standard textbook expositions of the discrete type adverse selection model (see for example Laffont and Martimort (2002)) it is usual to argue on the basis of the geometry of the agents' indifference curves in contract space (Φ-intervals, in our terminology), relying on the similarity to linear geometry. Our framework can be considered a sytematic generalisation of this approach.

As in standard convex analysis, there are two ways to define convexity of a function: internal and external. The function $\tau(x)$ is *internally Φ-convex* if the epigraph $\operatorname{epi}\tau = \{(x,t) : t \geq \tau(x)\}$ is an internally Φ-convex subset of $\mathbb{R} \times \mathbb{R}$. It is *externally Φ-convex* if it is the upper envelope of elementary functions in Φ. That is, in the classical case, if its epigraph is the intersection of the affine half spaces containing it. These two concepts are of course not the same.

The existing literature on abstract convexity focuses solely on outer convexity; the internal concept appears not to have been studied. Rather awkwardly, and in conflict with classical convex analysis, in this literature a function $\tau(x)$ is defined to be Φ-convex if it is the upper envelope of elementary functions in Φ.[9] We will attempt to avoid this confusion by calling outer convex functions envelope functions. If we occasionally use the term Φ-convex without qualification it will always mean internally Φ-convex, as in standard convex analysis. We note that if τ is a proper function bounded below by an elementary function then $\tilde{\tau} = \sup\{\phi \in \Phi : \phi \leq \tau\}$ is an envelope function: it is the greatest envelope function that minorises τ, and τ is an envelope function if and only if $\tau = \tilde{\tau}$. We call $\tilde{\tau}$ the Φ envelope of τ.

If this generalized convexity concept is to be useful, the minimal requirement is that the elementary functions Φ indeed be (internally) convex. We will say that $v(x,\theta)$ has *monotone differences* if, for all $\theta, \theta' \in \Theta$, the function $v(\cdot,\theta') - v(\cdot,\theta)$ is monotonic. That is either, for all $x' > x$, $v(x',\theta') - v(x,\theta') \geq v(x',\theta) - v(x,\theta)$ or, for all $x' > x$, $v(x',\theta') - v(x,\theta') \leq$

[9] Rockafellar and Wets sensibly avoid this terminology.

$v(x',\theta) - v(x,\theta)$. This is implied if $v(x,\theta)$ is supermodular or if $v(x,\theta)$ is submodular, or if there is a mixture of cases. It is equivalent to the requirement that, after an appropriate relabelling of types, $v(x,\theta)$ is supermodular.

Lemma 1 *The elementary functions Φ are internally Φ-convex if and only if $v(x,\theta)$ has monotone differences.*

Proof. The following are equivalent

1. $v(x,\theta)$ does not have monotonic differences
2. There exist $\theta, \theta' \in \Theta$ and $x < x' < x'' \in X$ such that $h(x) = v(x,\theta') - v(x,\theta)$ is strictly decreasing over one of the intervals $\{x, x'\}, \{x', x''\}$ and strictly increasing over the other
3. There exist θ, $\theta' \in \Theta$ and $x < x' < x'' \in X$ such that $h(x) = v(x,\theta') - v(x,\theta)$ is strictly decreasing over the interval $\{x, x'\}$ and strictly increasing over $\{x', x''\}$ (just switch θ, θ' if necessary)
4. There exist θ, $\theta' \in \Theta$, $x < x' < x'' \in X$ and a constant c such that $h(x) - c > 0$, $h(x') - c < 0, h(x'') - c > 0$
5. There exist elementary functions $f(x) = v(x,\theta')$ and $g(x) = v(x,\theta) + c$ such that $f(x) > g(x)$, $f(x') < g(x')$, $f(x'') > g(x'')$
6. There exists an elementary function $g(x)$ that is not internally convex.

■

Since, after relabelling types, monotone differences is equivalent to supermodularity, we shall in the rest of the paper impose the condition in this way. In fact, in order to get clean results, we shall strengthen this assumption and impose strict supermodularity.[10] A related, slightly stronger condition, which we will call virtual single crossing, will also be introduced below.

3.1 Φ-envelope functions

The key step in exploiting the power of the classical Fenchel transform is to characterise the functions that can arise as the envelope of a family of affine functions. The main result in this section will be an analogous characterization of the Φ-envelope functions under a strict supermodularity assumption. We also establish properties of a generalised Fenchel transform in this context. Since an envelope function is supported by a family of elementary functions, its growth rate is constrained by the growth rate of the elementary functions. To capture this the following definition will be useful.

Condition 1 (growth condition) *The function $f(z)$ satisfies the Φ growth condition if, for all $x < x'$ such that at least one of $f(x)$, $f(x')$ is finite*

$$\inf_{\theta} \left(v(x', \theta) - v(x, \theta) \right) \leq f(x') - f(x) \leq \sup_{\theta} \left(v(x', \theta) - v(x, \theta) \right).$$

Note that, under supermodularity, the elementary functions in Φ satisfy 1, and that this condition is preserved when passing to upper envelopes, as are both internal and external Φ-convexity. The key step in proving the main result is the following Lemma.

Lemma 2 *Assume that $v(x, \theta)$ is continuous and strictly supermodular, and that $\tau(x)$ is proper, internally Φ-convex and lower semicontinuous. Then the correspondence $\xi(\theta) = \operatorname{argmax}_{x \in X} v(x, \theta) - \tau(x)$ is maximal monotone.*

The proof is in the Appendix.

[10] If we allow nonstrict supermodularity the results become much messier as pooling can arise not only because of kinks in the contract but because the agent's marginal utility is not responsive to type in some regions. Since this behaviour is driven by the way that the agents' preferences are parametrised, and this can be arbitrarily complicated, we avoid these cases.

Theorem 2 *Assume that $v(x, \theta)$ is continuous and strictly supermodular. The following are equivalent*

1. *$\tau(x)$ is a Φ-envelope function. That is, it is externally Φ-convex;*

2. *$\tau(x)$ is proper, lower semicontinuous, internally Φ-convex and satisfies the growth condition 1;*

3. *$\tau(x)$ is proper, internally Φ-convex and satisfies the growth condition 1, and the correspondence ξ is maximal monotone.*

Proof. Condition (1) implies (2) because the elementary functions have these properties, and they are preserved in passing to the envelope by taking the intersection of epigraphs. Condition (2) implies (3) by Lemma 2. It remains to be shown that condition (3) implies (1).

Let $\tilde{\tau}$ be the Φ-convexification of τ (the supremum of the elementary minorants τ). This clearly exists (its epigraph is the intersection of the epigraphs of the primitive minorants) and $\tau \geq \tilde{\tau}$. We must show that $\tau = \tilde{\tau}$. Now $\tau(x) = \tilde{\tau}(x)$ if and only if $x \in \operatorname{im} \xi$, so we must show that $\operatorname{im} \xi = X$. Since ξ is maximal monotone $\operatorname{im} \xi$ is a connected interval, so if there exists a point such that $\tilde{\tau}(x) \neq \tau(x)$ then it must lie either above of below the interval $\operatorname{im} \xi$. We consider the first alternative, and assume that there exists $x' \in X$ such that $x' > \operatorname{im} \xi$ and $\tilde{\tau}(x') < \tau(x')$. Choose t' such that $\tilde{\tau}(x') < t' < \tau(x')$ and consider the contract (x', t'). No agent will accept this contract since any elementary function supporting the contract must pass through or below $(x', \tilde{\tau}(x'))$. Thus, for all $\theta \in \Theta$ and $x \in \xi(\theta)$, $v(x, \theta) - \tau(x) > v(x', \theta) - t'$. This implies that $\tau(x) + v(x', \theta) - v(x, \theta)$ is uniformly bounded above by t', and this implies, by the growth condition, that $\tau(x') < \infty$. The growth condition applies here since $x \in \operatorname{im} \xi$, so $\tau(x)$ is finite). So we can write

$t' = \tau(x') - \delta$. The incentive compatibility condition can then be written

$$\begin{aligned}\tau(x') - \tau(x) &> \delta + v(x', \theta) - v(x, \theta) \\ &\geq \delta + \sup_{\theta} v(x', \theta) - v(x, \theta) \\ &> \sup_{\theta} v(x', \theta) - v(x, \theta)\end{aligned}$$

which contradicts the growth condition. A similar argument applies if $x > \operatorname{im}\xi$ and $\tau(x) < \infty$. ■

In understanding the role played by the growth condition the following example is instructive. Let $X = \mathbb{R}$, let $\Theta = [-1, 1]$, let $v(x, \theta) = \theta x$, and consider the function $\tau(x) = \frac{x^2}{2}$. The family of primitive convex functions is $\Phi = \{\theta x - t : \theta \in [-1, 1], t \in \mathbb{R}\}$. Then τ is not an envelope function. The greatest envelope function less than or equal to τ, the supremum of the minorants in Φ, is $\tilde{\tau}(x) = \begin{cases} \frac{x^2}{2}, & |x| \leq 1 \\ |x| - \frac{1}{2}, & |x| \geq 1 \end{cases}$.

The classical characterization of envelope functions is equivalent to a separating hyperplane theorem of convex analysis (Rockafellar 1970). In a similar manner, Theorem 2 can be reformulated as a non-convex separation theorem.

Corollary 2 *Assume that $\tau(z)$ is proper, lower semicontinuous, internally Φ-convex and satisfies the growth condition 1, and let $t \leq \tau(x) < \infty$. Then there exists an elementary function $\phi \in \Phi$ such that $t \leq \phi(x)$ and, for all z, $\phi(z) \leq \tau(z)$.*

The Fenchel transform now generalizes in a straight forward way to this abstract convexity setting (see Rockafellar and Wets (1998, Section 11.L)).[11]

[11] The only aspect of the classical duality that is not generalised here is the identification in the linear case of $\xi(\theta)$ and $\eta(x)$ with the subgradients $\partial\rho(\theta)$ and $\partial\tau(x)$ respectively, since the sugradient is a global concept that is not adapted to the non-convex environment studierd here. It would presumably be possible to use one of the many local nonsmooth generalisations of the subgradient.

Let

$$\tau^*(\theta) = \sup_{x\in X} v(x,\theta) - \tau(x) \tag{4}$$
$$\xi(\theta) = \operatorname*{argmax}_{x\in X} v(x,\theta) - \tau(x) \tag{5}$$
$$\tau^{**}(x) = \sup_{\theta\in\Theta} v(x,\theta) - \tau^*(\theta) \tag{6}$$

be the generalised Fenchel Φ-conjugate and bi-conjugate of τ . The following properties, which are standard in the classical case, generalize immediately to this context. Note in particular the identification of the Φ envelope $\tilde{\tau}$ of τ with the biconjugate τ^{**}.

Lemma 3 (abstract Fenchel transform) *Let $\tau^*(\theta)$ and $\tau^{**}(x)$ be defined by the equations above. Then*

1. $(\theta, t) \in \operatorname{epi} \tau^* \Leftrightarrow v(x,\theta) - t$ *is a minorant of* τ
2. $(x, t) \in \operatorname{epi} \tau \Rightarrow v(x,\theta) - t$ *is a minorant of* τ^*
3. $\tilde{\tau} = \tau^{**}$
4. $\tau^{**} = \tau$ *if and only if* τ *is an envelope function*
5. $\tau(x) + \tau^*(\theta) \geq v(x,\theta)$ *(the Fenchel inequality)*
6. *The Fenchel inequality holds with equality if and only if* $x \in \xi(\theta)$ *, if and only if* $\theta \in \eta(x)$ *.*

Proof. These results are straight forward: see Rockafellar and Wets (1998, Sections 11.A, 11.L), and are elementary apart from (3) , which relies on the separation theorem proved above. To prove (3) we argue as follows. If ϕ is an elementary function such that $\phi \leq \tau$ then $\phi = \phi^{**} \leq \tau^{**}$, since the Fenchel transform is order reversing, and biconjugation order preserving. It thus follows that $\tilde{\tau} \leq \tau^{**}$. But τ^{**} is Φ convex, lower semicontinuous and satisfies the growth condition, since it inherits these properties from the elementary functions of which it is a supremum. By the separation theorem, $\tau^{**} \leq \tilde{\tau}$. ■

In the classical case the supremum in Equation 4 is actually attained as a maximum, except perhaps at the endpoints of $\operatorname{dom} \tau$. To establish the analogue of this important property in our context it is necessary to introduce a refinement of the single crossing property.[12] In setting this up it is useful to establish some terminology. The elementary function $\psi(x)$ supports a function $\tau(x)$ if $\psi(x) \leq \tau(x)$ and, for some $x_0, \psi(x_0) = \tau(x_0)$. That is, $\psi(x)$ minorizes $\tau(x)$ and touches it at some point. We will say that $\psi(x)$ *virtually* supports $\tau(x)$ if $\psi(x) < \tau(x)$ and $\inf_x \tau(x) - \psi(x) = 0$. Strict single crossing implies that if $\psi(x), \psi'(x)$ are different elementary functions then it is not possible that $\psi(x)$ support $\psi'(x)$ at an interior point of X. It does not however exclude the possibility that one elementary function might virtually support another without their being identical. That is to say, they might meet asymptotically at infinity (consider for example the family of affine functions on $\{x \in \mathbb{R} : x > 0\}$). This is the behaviour that we want to exclude. The following condition, which might be considered a kind of strict single crossing at infinity condition, is useful for this purpose.

Condition 2 (virtual single crossing) *The family Φ of elementary functions satisfies the virtual single crossing condition on Θ if $\psi(x) = v(x, \theta) - t$ virtually supports $\psi'(x) = v(x, \theta') - t'$ only if $\theta = \theta'$ and $t = t'$.*

The virtual single crossing condition is guaranteed if Θ is compact, or if the graphs of elementary functions of different types must always cross (as in the classical case on $\mathbb{R}$, where the elementary functions are the affines).

Theorem 3 *A strictly supermodular family Φ of elementary functions satisfies the virtual single crossing condition if and only if every envelope function*

[12] Without this refinement we can show that the maximum is attained everywhere except on a possibly nondegenerate interval at each end of $\operatorname{dom} \tau$.

τ is supported by its elementary minorants at every point of its domain $\operatorname{dom}\tau$ *except, perhaps, at the boundary.*

For the proof, see the Appendix. There are two examples that are informative in understanding this result and its proof. In both examples we set $v(x,\theta) = \theta x$. The first is $\tau(x) = \begin{cases} -\sqrt{x}, & x \geq 0 \\ \infty, & x < 0 \end{cases}$, for $x \in \mathbb{R}$, with $\tau^*(\theta) = \begin{cases} \infty, & \theta \geq 0 \\ -\frac{1}{4\theta}, & \theta < 0 \end{cases}$, $\xi(\theta) = \begin{cases} \{\}, & \theta \geq 0 \\ \left\{\frac{1}{4\theta^2}\right\}, & \theta < 0 \end{cases}$,for $\theta \in \mathbb{R}$. Here $\operatorname{im}\xi = (0,\infty) \subsetneq \operatorname{dom}\tau = [0,\infty)$. $\tau(x)$ is supported at every point of its domain except for the isolated point 0. At 0 any supporting hyperplane would need to have infinite slope. In the second example $X = \mathbb{R}$, and $\Theta = (0,\infty)$. We set $\tau(x) = \begin{cases} x^2, & x \geq 0 \\ 0, & x < 0 \end{cases}$. Then $\tau(x)$ is not supported at any point $x \in (-\infty, 0]$. The elementary functions in this example do not satisfy Condition 2.

3.2 Regularity Properties, and an Envelope Theorem.

We can now use Φ-convexity arguments to establish some important regularity properties of envelope functions. These generalize standard properties of convex functions (see, for example, Rockafellar (1970, 13.3.3), Bonnans and Shapiro (2000, 2.3.5)). They will allow us to prove a general envelope theorem that works in our environment. First, note the following geometric Lemma.

Lemma 4 (chord lemma) *Assume that the elementary functions Φ are strictly supermodular and satisfy the virtual single crossing condition. Assume that $v(x,\theta)$ is continuous and that $\tau(x)$ is a Φ-envelope function. Consider points $x' \in \xi(\theta')$, $x'' \in \xi(\theta'')$ with $x' < x''$.*

1. *There exists a Φ-interval I, the graph of an elementary function $f(z)$ restricted to $[x', x'']$, joining $(x', \tau(x'))$ and $(x'', \tau(x''))$ in* $\operatorname{epi}\tau$.

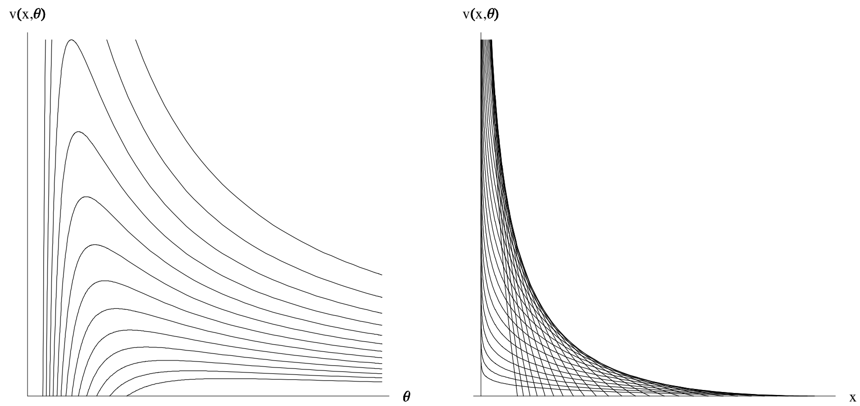

Figure 2: Envelope Example: $v(x,\theta) = -\frac{\log x}{\theta} - \frac{1}{\theta^2}$.

2. *$f(z) \geq \tau(z)$ if $x' \leq z \leq x''$, and $f(z) \leq \tau(z)$ if $z \leq x'$ or $z \geq x''$*

Proof. To check the existence of such an interval, consider the elementary function $\phi(z,\theta) = \tau(x') + v(z,\theta) - v(x',\theta)$. For any θ we have $\phi(x',\theta) = \tau(x')$. The quantity $\phi(x'',\theta)$ is continuous in θ. By incentive compatibility $\phi(x'',\theta') \leq \tau(x'')$ and $\phi(x'',\theta'') \geq \tau(x'')$, so for some intermediate θ with $\theta' \leq \theta \leq \theta''$ we must have $\phi(x'',\theta) = \tau(x'')$. Choosing this θ, we set $f(z) = \phi(z,\theta)$. This proves 1 and the first part of 2.

To prove the second part of 2, consider the case $z > x''$. It is sufficient to assume that $\tau(z) < \infty$, so $z \in \operatorname{dom}\tau$. We consider first the case that $z \in \operatorname{im}\xi$ so $z \in \xi(\theta''')$ for some θ'''. We extend the interval I by enlarging its domain of definition to $[x', z]$. By part 1, there exists a Φ-interval J joining $(x', \tau(x'))$ and $(z, \tau(z))$ in $\operatorname{epi}\tau$. The intervals I and J cross at $(x', \tau(x'))$. J is above I at x'' since J lies in $\operatorname{epi}\tau$ over the whole interval $[x', z]$ and I crosses the graph of τ at x''. If $f(z) > \tau(z)$ then J would be below I at z, contradicting single crossing.

It remains to consider the case that $z \in \operatorname{dom}\tau \setminus \operatorname{im}\xi$. By Theorem 3, and the standing assumptions, $\operatorname{im}\xi$ is dense in $\operatorname{dom}\tau$ so the result follows from continuity of τ on $\operatorname{dom}\tau$.[13] $\blacksquare$

Lemma 5 *Assume that the elementary functions* Φ *are strictly supermodular and satisfy the virtual single crossing condition. Assume that* $v(x,\theta)$ *is locally Lipschitz continuous in* x *for each* θ *and that* $\tau(x)$ *is a* Φ*-envelope function. Then* $\tau(x)$ *is locally Lipshitz , and hence continuous and absolutely continuous, at interior points of its domain.*

Proof. Since $\operatorname{im}\xi$ is connected and dense in $\operatorname{dom}\tau$, we can consider $x \in \xi(\theta)$ in the interior of $\operatorname{im}\xi$, and choose $x' \in \xi(\theta')$, $x'' \in \xi(\theta'')$ with $x' < x < x''$. By the Chord Lemma 4 there exists an elementary function ϕ' passing through $(x', \tau(x'))$ and $(x, \tau(x))$, and elementary function ϕ'' passing through $(x, \tau(x))$ and $(x'', \tau(x''))$, and an elementary function ϕ supporting τ at x. Then τ is sandwiched between the locally Lipshitz functions ϕ and $\max[\phi', \phi'']$ on the interval $[x', x'']$. $\blacksquare$

Theorem 4 (supermodular envelope theorem) *Assume that the elementary functions* Φ *are strictly supermodular and satisfy the virtual single crossing condition,and assume that* $v(x,\theta)$ *is continuous and differentiable in* x*. Let* $\tau(x)$ *be a* Φ *envelope function. That is,*

$$\tau(x) = \sup_{\theta\in\Theta} f(x,\theta)$$

$$\eta(x) = \operatorname*{argmax}_{\theta\in\Theta} f(x,\theta)$$

[13]For the result that a unidimensional lsc convex function is continuous on its domain see Rockafellar and Wets (1998, 2.37). We verify that this classical result continues to hold in our context. We already have lower semicontinuity, since τ is an envelope function. To show upper semicontinuity, assume that $z < \operatorname{im}\xi$, choose $\alpha > \tau(z)$, and choose a point $(\theta, x) \in \xi$. Let ϕ be the elementary function of type θ supporting τ at x. We shift this function up so that it passes through the point $(z, \tau(z))$ and check that the graph of this function lies in $\operatorname{epi}\tau$ on the interval $[x, z]$. We now have τ bounded above on this interval by a continuous function passing through (z, α). Since α is chosen arbitrarily $\geq \tau(z)$ we get upper semicontinuity.

where, for each $\theta \in \Theta$, $f(x,\theta) \in \Phi$. Then $\tau(x)$ is differentiable almost everywhere, $\tau'(x) = f_x(x, \theta(x))$ almost everywhere, and

$$\tau(x) = \tau(x_0) + \int_{x_0}^{x} f_x(x, \theta(x))\, dx$$

where $\theta(x)$ is any selection from $\eta(x)$ and x_0 is an interior point of $\operatorname{dom}\tau$.

Proof. Since $f(x,\theta) \in \Phi$ we can write $f(x,\theta) = v(x,\theta) - t(\theta)$ for some function $t(\theta)$ on Θ. That is, $\tau(x)$ is the Fenchel transform of $t(\theta)$. By fenchel duality (Lemma 3) we can take $t(\theta) = \tau^*(\theta)$. By Lemma 5 τ is absolutely continuous so it is differentiable almost everywhere and $\tau(x) = \tau(x_0) + \int_{x_0}^{x} \tau'(x)\, dx$. Let x be a point of differentiability of τ in the interior of $\operatorname{dom}\tau$. By Theorem 3 τ is supported at x by the elementary function $\phi(z) = v(z,\theta) - \tau^*(\theta)$ for some $\theta \in \eta(x)$. Since ϕ supports τ at x, and both functions are differentiable at this point, it follows that $\tau'(x) = \phi'(x) = v_x(x, \theta(x)) = f_x(x, \theta(x))$. ■

Notice that it is not sufficient to assume that $v(x,\theta)$ is differentiable almost everywhere (or even absolutely continuous) with respect to x, as is shown by the example $f(x,\theta) = -\theta^2 + 2\theta x - |x - \theta|$.

This result is related to, but different from, the Envelope Theorem of Milgrom and Segal (2002) who study the general parametric optimisation problem

$$\tau(x) = \sup_{\theta \in \Theta} f(x,\theta).$$

Milgrom and Segal assume that $f(x,\theta)$ is absolutely continuous and everywhere differentiable in x, and require the side condition that the derivative $f_x(x,\theta)$ be uniformly integrable, something that is not necessarily easy to check if the type space or the action space is not compact.[14] They show

[14] Milgrom and Segal note in a footnote to their paper that their uniform integrability condition will hold in a quasilinear environment with single crossing provided that both the type space and the action space are both compact intervals.

that the value function $\tau(x)$ is absolutely continuous and derive an envelope relationship between $\tau(x)$ and the parametric family $f(x,\theta)$.

In the quasilinear environment (which of course is a more restrictive environment than that studied by Milgrom and Segal) things are somewhat simpler. Instead of their integral side condition it is necessary only to check single crossing and virtual single crossing, order conditions that are often automatic in applications. We also get a stronger result: $\tau(x)$ is not just absolutely continuous but locally Lipshitz (this will be useful below, in Theorem 8). Consider the example

$$\tau(x) = \sup_{\theta\in\Theta} -\frac{\log x}{\theta} - \frac{1}{\theta^2}$$

with $X = \Theta = (0,1]$, which satisfies the conditions of Theorem 4, but which does not satisfy the Milgrom Segal condition unless both the type space and the action space are restricted to compact intervals away from 0. Figure 2 shows that this example is not in any way pathological.

4 Quasilinear Agency

We are now in a position to return to the general agency problem that we started with: an agent of type θ who derives utility $v(x,\theta)-t$ from purchasing the allocation x by transfer t, confronts the problem

$$\begin{aligned}\rho(\theta) &= \sup_{x\in X} v(x,\theta) - \tau(x) \\ \xi(\theta) &= \operatorname*{argmax}_{x\in X} v(x,\theta) - \tau(x).\end{aligned}$$

We assume that $v(x,\theta)$ is strictly supermodular, is continuously differentiable, and satisfies the virtual single crossing condition.

Consider first the information rent $\rho(\theta)$ that the agent might receive. How do we characterise the rent functions $\rho(\theta)$ that can be implemented by

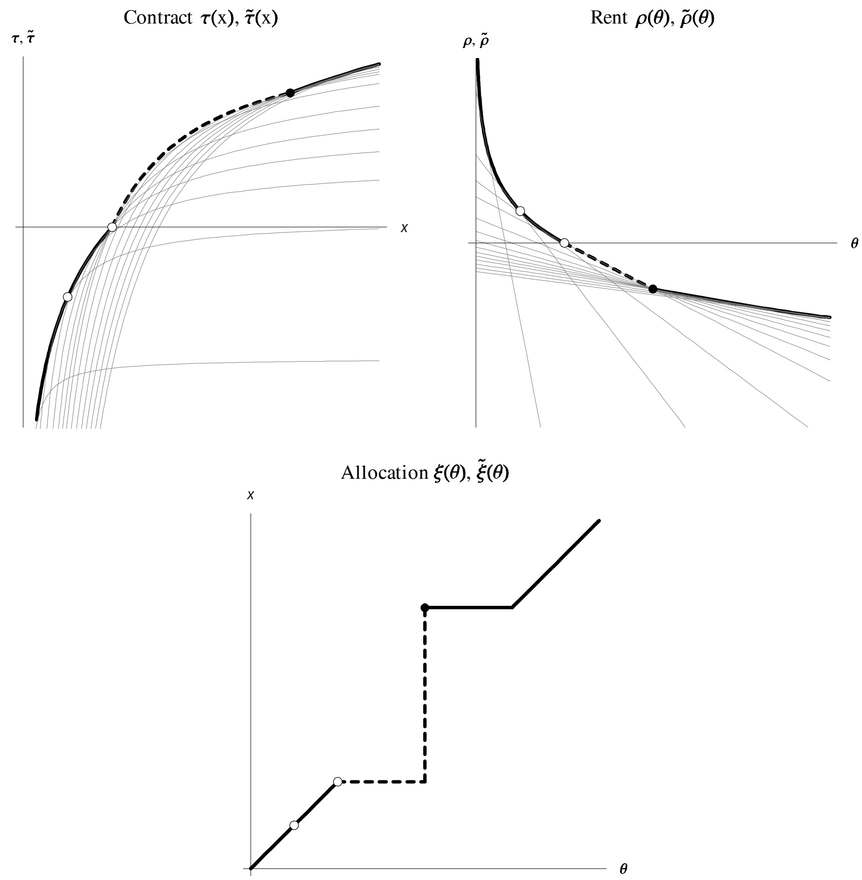

Figure 3: Implementation of ξ when the valuation is $v(\theta, x) = \frac{x-\theta}{x}$.

a contract? We need to ensure two things: that $\rho(\theta)$ is an envelope function, and that for each type θ there actually exists an allocation that supports ρ at θ. That is, the supremum in the agent's decision problem is attained as a maximum. Theorems 2 and 3 address these issues. We summarise.

Theorem 5 *$\rho(\theta)$ is an envelope function if and only if it is proper, lower semicontinuous, internally Φ-convex and satisfies the growth condition 1. In that case $\rho(\theta)$ is supported by an allocation $x \in \xi(\theta)$ at every $\theta \in \operatorname{dom} \tau$, except possibly at the endpoints of the domain.*

Now that we know that ρ is an envelope function, implemented by some contract τ, we observe that ρ is the generalised Fenchel dual of τ, so in fact we can use $\tau = \rho^*$ as the implementing contract.

Theorem 6 *Let ρ be an envelope function implemented by $\tau = \rho^*$. Then ρ and τ are locally Lipshitz, absolutely continuous at interior points of their domains, and satisfy*

$$\begin{aligned}\rho(\theta) &= \rho(\theta_0) + \int_{\theta_0}^{\theta} v_\theta(x(\theta), \theta)\, d\theta \\ \tau(x) &= \tau(x_0) + \int_{x_0}^{x} v_x(x, \theta(x))\, dx\end{aligned}$$

where θ_0, x_0 are interior points of $\operatorname{dom}\rho, \operatorname{dom}\tau$ respectively, $x(\theta)$ is a selection from $\xi(\theta)$ and $\theta(x)$ is a selection from $\eta(x) = \xi^{-1}(x)$.

Finally, just as in the classical linear case, we can characterise the implementable allocations.

Theorem 7 *Let $\xi \subset \Theta \times X$ be an allocation correspondence. Then ξ is implementable in the quasilinear contracting environment if and only if*

1. *ξ is monotone;*
2. *$\xi(\theta)$ is convex for all $\theta \in \Theta$;*
3. *$\eta(x)$ is closed and convex for all $x \in X$; and, at any gap $\xi_- < \xi_+$,*
4. *either $\eta(x_-) = \emptyset$, or $\eta(x_+) = \emptyset$ (both may be empty);*
5. *if $\eta(x_+) \neq \emptyset$ then $\theta_+ \in \eta(x_+)$; if the gap is at the bottom then $(-\infty, \theta_+] \subset \eta(x_+)$;*
6. *if $\eta(x_-) \neq \emptyset$ then $\theta_- \in \eta(x_-)$; if the gap is at the top then $[\theta_-, \infty) \subset \eta(x_+)$.*

The proof is exactly the same as in Section 2.1. Just as in the linear case, ξ is maximal monotone if and only if the implementing contract τ is an envelope function. Corollary 1 and Proposition 1 also apply without change.

To illustrate this result consider the following example, which is illustrated in Figure 3. The allocation is the same as in Figure 1, but the valuation $v(\theta, x) = \frac{x-\theta}{x}$ is nonlinear. The contract that supports ξ is $\tau(x) = \log x$ if $0 < x < \frac{1}{2}$ or $\frac{1}{2} < x < 1$, and $\tau(x) = \frac{4}{3} - \log 3 + \log x$ if $x \geq 3$. The information rent is $\rho(\theta) = -\log\theta$ if $0 < \theta < \frac{1}{2}$, $\frac{1}{2} < \theta < 1$, and $\rho(\theta) = -\frac{4}{3} + \log 3 - \log\theta$ if $\theta \geq 3$. The contract that supports the maximal monotone extension $\tilde{\xi}$ is $\tilde{\tau}(x) = \log x$ if $0 < x < 1$, $\tilde{\tau}(x) = 2 - \frac{2}{x}$ if $1 < x \leq 3$, and $\tilde{\tau}(x) = \frac{4}{3} - \log 3 + \log x$ if $x \geq 3$. The information rent is $\tilde{\rho}(\theta) = -\log\theta$ if $0 < \theta < 1$, $\tilde{\rho}(\theta) = 1 - \theta$ if $1 < \theta \leq 2$, $\tilde{\rho}(\theta) = -\frac{1}{3} - \frac{\theta}{3}$ if $2 < \theta \leq 3$, and $\tilde{\rho}(\theta) = -\frac{4}{3} + \log 3 - \log\theta$ if $\theta \geq 3$.

4.1 Application: menus of simple contracts

It is common in applications to assume that the principal offers a menu of linear contracts, but this is a very restrictive assumption (Rogerson 1987). Consider for example an agency model with a smooth, strictly convex contract $\tau(x)$ and a linear contract $l(x)$ that supports $\tau(x)$ at x_0. That is, $l(x_0) = \tau(x_0)$ and $l'(x_0) = \tau'(x_0) = \theta_0$. Then every point on the graph of $l(x)$, apart from $(x_0, l(x_0))$ lies strictly outside the contract set $\operatorname{epi}\tau$. If this contract is offered as part of a menu then any agent of type $\theta \neq \theta_0$ will misrepresent their type as θ_0 and select a point on $l(x)$. A similar case can be made using the nonlinear example discussed in the previous Section.

Let us define a simple contract as one that is continuous and piecewise linear with a single kink: $\tau(x) = \begin{cases} \tau_0 + \alpha(x - x_0), & x \leq x_0 \\ \tau_0 + \beta(x - x_0), & x \geq x_0. \end{cases}$ That is to say, τ is characterized by a target x_0, a linear penalty for falling short of x_0, and a linear reward for exceeding x_0.

Theorem 8 *Assume that $v(x,\theta)$ is continuous, Lipshitz continuous in x, and strictly supermodular, and that it satisfies the virtual single crossing condition. Let ξ be an allocation correspondence that is supported by an incentive compatible contract τ. Then ξ is supported by a menu of simple contracts.*

Proof. Without loss of generality one can assume that τ is an envelope contract, since any simple contract that supports $\tilde{\tau}$ will support τ. Let $(\theta_0, x_0) \in \xi$. By the Fenchel inequality $\tau(x) \leq v(x,\theta_0) - \rho(\theta_0)$, with equality when $x = x_0$. But $v(x,\theta_0)$ is Lipschitz, so it is supported from above by a simple contract. ■

It is worth noting that this result is different in nature to properties established in, for example, the elementary part of Theorem 2. Envelope functions are by definition constructed by approximation from below, and properties like lower semicontinuity and Φ-convexity that are stable under approximation from below are more or less automatic. Theorem 8, which relates to approximation from above, is more delicate.

The Lipshitz hypothesis means in essence that there are constant or decreasing returns to scale in the value $v(x,\theta)$ to type θ of the allocation x.

5 Conclusion

If the agent's valuation is linear, then the standard adverse selection contracting problem is isomorphic to the producer's problem in elementary price theory, with price replaced by the agent's type. Standard duality results of price theory, which are most conveniently expressed using the Fenchel transform, apply more or less directly. In particular, not only the information rent but also the contract can be expressed as an integral of marginal valuations. If the environment is quasilinear, and the agent's valuation is strictly supermodular, then an abstract convexity and an abstract Fenchel transform can be developed which allow these results to be generalised almost in their entirety to the quasilinear contracting problem. The supermodular quasilinear contracting problem is one of the canonical models of microeconomics, with applications ranging from optimal regulation to auction theory.

This framework can be used to establish an envelope theorem, a complete duality theory between the information rent and the contract, and a characterization of implementable allocations. Supermodularity plays a key role in two respects. Firstly, it creates a link between incentive theory and an abstract convexity that is rich enough to capture both inner and outer concepts of convexity and the relations between them. This structure allows standard techniques of convex optimisation to be applied in this generalised setting. Secondly, it implies that the allocation correspondence is maximal monotone. Maximal monotone correspondences have strong properties that drive the results.

The framework developed in this paper has potential for applications that go beyond those presented here. In this paper we focus on the duality structure inherent in the agent's decision problem, and draw out some implications for the structure of implementable contracts. There is no discussion of optimal contracts. It seems likely that there are useful things to be said in this direction, building in particular on implications of maximal monotonicity that go beyond the techniques used here. The other interesting direction is multidimensional type spaces. While the details of particular arguments in this paper draw heavily on the one dimensionality assumption, the general framework of Fenchel duality on which we draw is inherently multidimensional. It is natural to conjecture that there may be useful generalizations to this setting.

A Appendix: Proofs

Proof of Theorem 1. It will be convenient throughout the proof to write $K = \operatorname{im} \xi = \operatorname{dom} \eta$ and $L = \operatorname{im} \eta = \operatorname{dom} \xi$. We confirm first of all the necessity of the conditions. Assume that $\xi(\theta) = \operatorname{argmax}_{x \in X} \theta x - \tau(x)$ is implementable by a contract $\tau(x)$. Using the normalisation imposed above, we have $\tau(x) = \tilde{\tau}(x)$ on $K = \operatorname{dom} \tau \subset \operatorname{dom} \tilde{\tau}$ and $\tau(x) = \infty$ outside K. Since $\tilde{\tau}$ is lower semicontinuous it is continuous on its domain (Rockafellar and Wets (1998, 2.35)), so τ is continuous on K. Monotonicity of ξ follows directly from the single crossing property, as does convexity of the sets $\xi(\theta)$ and $\eta(x)$. That $\eta(x)$ is closed follows from continuity of the valuation θx. If $\theta_i \in \eta(x)$ then $\theta_i(z - x) + \tau(x)$, the type θ_i indifference curve through $(x, \tau(x))$, supports τ at x. If $\theta_i \to \theta$ then by continuity $\theta(z - x) + \tau(x)$ also supports τ at x, so $\theta_i \in \eta(x)$.

To show (5), assume that $(\theta, x_+) \in \xi_+$. If $(\theta_+, x_+) \in \xi_+$ then we are done; if not, then we can approach (θ_+, x_+) from within ξ_+ along the horizontal line through (θ, x_+). Thus there exist $\theta_i \in \eta(x_+)$ with $(\theta_+, x_+) \leq (\theta_i, x_+) \leq (\theta, x_+)$ and $\theta_i \to \theta_+$, and the result follows since $\eta(x_+)$ is closed. If the gap is at the bottom and $\theta \leq \theta_+$ then by single crossing θ will also choose x_+. The proof of (6) is similar. To prove (4), assume the contrary. Then it must be the case that $(\theta_-, x_-) \in \xi$ and $(\theta_+, x_+) \in \xi$. Assume for the moment that $\theta_- < \theta_+$ and let θ be such that $\theta_- < \theta < \theta_+$. Then either $\theta x_- - \tau(x_-) \geq \theta x_+ - \tau(x_+)$ or $\theta x_- - \tau(x_-) \leq \theta x_+ - \tau(x_+)$; that is, either θ weakly prefers x_- to x_+, or the converse. Assume the former. Then, by single crossing, θ will not strictly prefer any $z \geq x_+$ or $z < x_-$ to x_-. Nor will θ prefer any x with $x_- < x < x_+$ since in this range $\tau(x) = \infty$. Thus θ supports x_-. But this contradicts the maximality of θ_-. So it must be the case that $\theta_- = \theta_+ = \theta$. Hence $\xi(\theta) = \{x_-, x_+\}$. This, however, contradicts the convexity of $\xi(\theta)$.

Conversely, let ξ be a correspondence with the properties (1) to (6). If ξ is in fact maximal monotone then the result follows from Fenchel duality. Choose $x_0 \in K$ and consider the contract $\tilde{\tau}(x) = \tilde{\tau}(x_0) + \int_{x_0}^{x} \theta(x)\, dx$ (this is unique, up to the additive constant $\tilde{\tau}(x_0)$). Then $\tilde{\tau}$ implements a maximal monotone correspondence $\tilde{\xi}$ that contains ξ. Since ξ is maximal monotone $\xi = \tilde{\xi}$, so $\tilde{\tau}$ implements ξ.

If ξ is not maximal monotone then it can be enlarged to a maximal monotone correspondence $\bar{\xi}$, implemented as in the previous paragraph by a contract $\bar{\tau}$. We will show that $\bar{\xi}$ and $\bar{\tau}$ can be chosen so that $\tau = \bar{\tau}|_K$ implements ξ. To enlarge ξ it is sufficient to show how to fill in any gap $\xi_- < \xi_+$. It is more convenient to specify the inverse correspondence $\bar{\eta}$ rather than $\bar{\xi}$. If the gap is at the top, so $\xi_+ = \emptyset$, then we set $\eta(x_-) = [\theta_-, \infty)$. Similarly, if the gap is at the bottom, so $\xi_- = \emptyset$, then we set $\eta(x_+) = (-\infty, \theta_+]$. So we need only consider the case where the gap is neither at the top nor the bottom. By monotonicity it must be that $\theta_- \leq \theta_+$. Choose θ' such that $\theta_- \leq \theta' \leq \theta_+$ and set $\eta(x_-) = [\theta_-, \theta']$, $\eta(x_+) = [\theta', \theta_+]$, and $\eta(x) = \theta'$ for $x_- < x < x_+$. That is, types θ with $\theta_- \leq \theta \leq \theta'$ pool at x_-, type θ' mixes over the interval $[x_-, x_+]$, and types θ with $\theta' \leq \theta \leq \theta_+$ pool at x_+. It is easy to check that $\bar{\xi}$, so constructed is maximal monotone, that $\xi = \bar{\xi}|_L = \left\{(\theta, x) \in \bar{\xi} : \theta \in L\right\}$, and that $\bar{\tau}(x)$ is affine on $[x_-, x_+]$ since it is the integral of a constant.

Let $\tau = \bar{\tau}|_K$. We show that τ implements ξ; that is, $\xi(\theta) = \operatorname{argmax}_x \theta x - \tau(x)$ for all $\theta \in \Theta$. Let $x \in \xi(\theta)$, so $(\theta, x) \in \xi \subset \bar{\xi}$. Since $\bar{\tau}$ implements $\bar{\xi}$, the θ indifference curve through $(x, \tau(x)) = (x, \bar{\tau}(x))$ supports $\bar{\tau}$ at x, so it also supports τ at x. Thus $\xi(\theta) \subset \operatorname{argmax}_x \theta x - \tau(x)$.

To show the converse we argue as follows. If $x \in \operatorname{argmax}_x \theta x - \tau(x)$ then θ chooses x from the menu τ. In particular, $\tau(x) < \infty$, so $x \in K$. We show that θ will in fact choose x from the larger menu $\bar{\tau}$. So assume that $x \in K$ and θ strictly prefers x' to x when choosing from $\bar{\tau}$. Since x is optimal in K, it follows that $x' \notin K$ so $(\theta, x') \notin \xi$. It must be the case that (θ, x') lies in a gap $\xi_- < (\theta, x') < \xi_+$. We assume that $x < x'$; the converse case is similar. If $\xi_+ = \emptyset$ then the argument is straight forward, so we assume that $\xi_+ \neq \emptyset$.

By the convexity of $\bar{\tau}$, it follows that when choosing from $\bar{\tau}$, θ strictly prefers any $x'' \in (x, x']$ to x. For assume the contrary: some such x'' is not strictly preferred to x. This means that $(x'', \bar{\tau}(x''))$ lies on or above the affine line of slope θ through $(x, \bar{\tau}(x)) = (x, \tau(x))$ (the indifference curve of type θ), and $(x', \bar{\tau}(x'))$ lies strictly below it. Let $\tilde{\theta} < \theta$ be the slope of the chord through $(x, \bar{\tau}(x))$ and $(x', \bar{\tau}(x'))$, and consider the affine of slope $\tilde{\theta}$ through $(x, \bar{\tau}(x))$. Then $(x, \bar{\tau}(x))$ and $(x', \bar{\tau}(x'))$ lie on this line but $(x'', \bar{\tau}(x''))$ lies strictly above it. This contradicts the convexity of $\operatorname{epi} \bar{\tau}$.

As a consequence, it must be the case that $x = x_-$. Now $\bar{\tau}(z)$ is affine on $[x_-, x_+]$; that is to say, it coincides with an indifference curve of type θ', as constructed above. So by single crossing, if θ strictly prefers $x' \in [x_-, x_+]$ to x_- then it must be that $x' = x_+$. So θ strictly prefers x_+ to $x = x_-$. But there exist points x_i in K arbitrarily close to x_+, so θ must strictly prefer some such x_i to x. This contradicts the assumption that $x \in \operatorname{argmax}_x \theta x - \tau(x)$.

That is to say, that θ strictly prefers x on K. ■

Proof of Lemma 2. We first check that the correspondence $\theta \mapsto \xi(\theta)$ is monotonic. Let $x \in \xi(\theta)$, $x' \in \xi(\theta')$, with $\theta' < \theta$. We must check that $x' \leq x$. By incentive compatibility $v(x,\theta) - \tau(x) \geq v(x',\theta) - \tau(x')$ and $v(x',\theta') - \tau(x') \geq v(x,\theta') - \tau(x)$. Adding these inequalities, $v(x',\theta') - v(x,\theta) - v(x',\theta') + v(x,\theta) \geq 0$. But this contradicts the supermodularity assumption unless $x' \leq x$.

Consider first the case where X is compact, and note that in this case the correspondence ξ is closed, non-empty compact valued, and upper hemicontinuous. This is, in essence, the Berge maximum theorem (Berge 1963), but since this is usually stated under an assumption of continuity, not semicontinuity, we outline the argument. It is immediate that ξ is non-empty compact valued since X is compact and τ is lower semicontinuous. Assume that ξ is not closed. Then there exist $x_i \to x$, $\theta_i \to \theta$, $x_i \in \xi(\theta_i)$, but $x \notin \xi(\theta)$. That is, there exists z such that $v(x,\theta) - \tau(x) < v(z,\theta) - \tau(z)$. Now, by the lower semicontinuity of τ and the continuity of v we have

$$\begin{aligned}\limsup v(x_i,\theta_i) - \tau(x_i) &\leq v(x,\theta) - \tau(x) \\ &< v(z,\theta) - \tau(z) \\ &= \lim v(z,\theta_i) - \tau(z)\end{aligned}$$

so eventually there exists i such that $v(x_i,\theta_i) - \tau(x_i) < v(z,\theta_i) - \tau(z)$, which contradicts the assumption that $x_i \in \xi(\theta_i)$. So the correspondence ξ is closed. Since X is compact it then follows by standard results (see Aliprantis and Border (2006, Lemma 14.12)) that ξ is upper hemicontinuous.

Since for all θ it is the case that $\xi(\theta) \neq \emptyset$, to show that ξ is maximal monotone it is sufficient to show that the image $\Xi = \operatorname{im}\xi$ of ξ is connected.[15] Let $a < b$ be in the range Ξ, and let x be such that $a < x < b$ but $x \notin \Xi$. Let $x_- = \sup\{z \in \Xi : z < x\}$, and let $x_+ = \inf\{z \in \Xi : z > x\}$. Now there exist x_i, θ_i such that $x_i \in \xi(\theta_i)$ and $x_i \downarrow x_+$, and hence θ_i is decreasing. Moving if necessary to a subsequence, we may assume that $\theta_i \downarrow \theta_+$ for some θ_+. By upper hemicontinuity, $x_+ \in \xi(\theta_+)$. Using again the fact that ξ is closed, we can choose θ_+ to be the minimal θ such that $x_+ \in \xi(\theta)$. Similarly, θ_- is the greatest θ such that $x_- \in \xi(\theta)$. If $\theta_- < \theta < \theta_+$ then by monotonicity $\xi(\theta) \subset$

[15] This follows, for example, from the construction of 1 dimensional maximal monotone correspondences by filling in the gaps in the graph of a monotone function, as outlined previously.

$[x_-, x_+]$ but it cannot take on any interior value. Thus $\xi(\theta) \subset \{x_-, x_+\}$. Arguing again by monotonicity, there exists a $\tilde{\theta}$ between θ_- and θ_+ such that $\xi(\theta) = \{x_-\}$ for $\theta < \tilde{\theta}$, and $\xi(\theta) = \{x_+\}$ for $\tilde{\theta} < \theta$. But this would contradict the minimality of θ_+ or the maximality of θ_-. Thus $\theta_- = \theta_+$, and $\xi(\theta_-) = \xi(\theta_+) = \{x_-, x_+\}$. So we have shown that there exists a θ such that $\xi(\theta) = \{x_-, x_+\}$. Since both x_- and x_+ are optimal for this type we must have $v(x_-, \theta) - \tau(x_-) = v(x_+, \theta) - \tau(x_+)$. Thus we have constructed a Φ-interval such that both endpoints lie in $\operatorname{epi}\tau$ but the interior points do not lie in $\operatorname{epi}\tau$. But this would contradict the Φ-convexity of $\operatorname{epi}\tau$, so it must be that the image of ξ is a connected interval.

Now consider the case where X is not necessarily compact. If ξ is not maximal monotone then there exists a point $(\theta, x) \in \Theta \times X$ such that $(\theta, x) \notin \xi$ yet (θ, x) is comparable to ξ in the product order on $\Theta \times X$; this means that for any $(\theta', x') \in \xi$ either $(\theta', x') \leq (\theta, x)$ or $(\theta', x') \geq (\theta, x)$. For then (θ, x) could be added to the correspondence ξ without destroying monotonicity. Let ϕ be the elementary function of type θ passing through $(x, \tau(x))$. If we can show that ϕ supports τ at x then we will have established by contradiction that ξ is maximal monotone.

Consider first the case where X has a maximal element, and x is this maximal element. We show first that τ is supported at x by an elementary function $\tilde{\phi}$ whose type we will denote $\tilde{\theta}$. Let $I = [x', x]$ be a non-degenerate closed interval in X with with upper bound x. We consider the problem restricted from X to I. Since I is compact we know by the paragraph above that τ is supported on I by an elementary function $\tilde{\phi}(x)$, whose type will be $\tilde{\theta}$. We need to check that $\tilde{\phi}$ supports τ on the whole of X, not just on I. Assume not. Then $\tilde{\phi}(x) = \tau(x)$, $\tilde{\phi}(z) \leq \tau(z)$ for $z \in I$, but $\tilde{\phi}(x'') > \tau(x'')$ for some $x'' \in X \diagdown I$. Consider Φ-interval formed by the graph of $\tilde{\phi}$ restricted to the interval $[x'', x]$. We have just shown that the endpoints of the interval lie in $\operatorname{epi}\tau$, so by internally Φ-convexity the whole Φ-interval lies in $\operatorname{epi}\tau$. In particular, $\tilde{\phi}(z) \geq \tau(z)$ for $z \in I$. But we have already shown the opposite inequality, so we conclude that $\tilde{\phi}(z) = \tau(z)$ for $z \in I$. We now consider the larger interval $J = [x'', x]$, and a point $\hat{x}$ in the interior of I. Appealing once again to the result that has been proven on compact intervals, there exists an elementary function $\hat{\phi}$ of type $\hat{\theta}$ that supports τ at $\hat{x}$ on J. But τ is equal to the elementary function $\tilde{\phi}$ on a neighbourhood of $\hat{x}$, so by strict single crossing $\hat{\phi} = \tilde{\phi}$ and $\hat{\theta} = \tilde{\theta}$. In particular, this means that $\tilde{\phi}$ must support τ on the whole of $J = [x'', x]$. But this contradicts the assumption that $\tilde{\phi}(x'') > \tau(x'')$. So we have proven that there is an elementary function of type $\tilde{\theta}$ that supports τ at x. That is, $\left(\tilde{\theta}, x\right) \in \xi$. Since by assumption $(\theta, x) \geq \xi$ it follows that $\theta \geq \tilde{\theta}$, so ϕ is steeper than $\tilde{\phi}$, and $\tilde{\phi}(x) = \phi(x)$. Thus ϕ supports τ at x. The case where X has a minimal element is similar.

We now turn to the general case. If $(\theta, x) \leq \xi$ or $(\theta, x) \geq \xi$ then by truncating X at x we are back to the case where the result has been proven. So we can decompose $\xi = \xi_1 \cup \xi_2$ into two disjoint, nonempty pieces such that $\xi_1 \leq (\theta, x) \leq \xi_2$. We split X into two corresponding pieces $X_1 = \{z \in X : z \leq x\}$ and $X_2 = \{z \in X : z \geq x\}$ as well. Arguing as above, there is an elementary function ϕ_1 of type θ_1 supporting τ at x on X_1, and an elementary function ϕ_2 of type θ_2 supporting τ at x on X_2. By monotonicity $\theta_1 \leq \theta_2$, so in fact ϕ_1 and ϕ_2 support τ on the whole of X. Thus $(\theta_1, x), (\theta_2, x) \in \xi_1, \xi_2$ respectively and $(\theta_1, x) \leq (\theta, x) \leq (\theta_2, x)$, so ϕ is sandwiched between ϕ_1 and ϕ_2. Thus ϕ supports τ at x on X. This contradiction establishes the result. ■

The following Lemma will be used in the proof of Theorem 3

Lemma 6 *Assume that τ is an envelope function. Let $\xi(\theta) = \operatorname{argmax}_{x \in X} v(x, \theta) - \tau(x)$ and $\phi_x(\theta) = v(x, \theta) - \tau(x)$. Then $\operatorname{im} \xi \subset \operatorname{dom} \tau$, $\operatorname{dom} \xi \subset \operatorname{dom} \tau^*$, and*

1. *$x \in \operatorname{im} \xi$ if and only if ϕ_x supports τ^**

2. *$x \in \operatorname{dom} \tau$ if and only if ϕ_x virtually supports τ^**

3. *if $\theta \in \operatorname{dom} \tau^*$ then for any $\varepsilon > 0$ there exists $\theta' \in \operatorname{dom} \xi$ and an elementary function $\phi_{x'}$ that supports τ^* at θ' such that $\phi_{x'}$ ε-supports τ^* at θ. That is, $\phi_{x'}(\theta) \leq \tau^*(\theta) \leq \phi_{x'}(\theta) + \varepsilon$.*

Proof. Part 1 follows because $x \in \operatorname{im} \xi$ if and only if there is equality in the Fenchel inequality. Part 2 follows because $\tau(x) = \sup_\theta v(x, \theta) - \tau^*(\theta)$ is the Fenchel transform of its dual τ^*. Assertion 3 is quite intuitive. In essence it says that given an ε support to τ^* at θ we can shift it upwards so that it touches the graph of τ^* somewhere, though not perhaps at θ. This observation will be useful in the proof below in determining the relative slopes of virtual supporting functions. To prove 3 we argue as follows. If $\theta \in \operatorname{dom} \xi$ then we are done; just set $\theta' = \theta$. So assume $\theta \notin \operatorname{dom} \xi$. By maximal monotonicity $\operatorname{dom} \xi$ is connected so either $\theta < \operatorname{dom} \xi$ or $\theta > \operatorname{dom} \xi$. We consider the latter. It must be the case that $\operatorname{im} \xi$ is a connected interval, unbounded above. For if it were bounded above, say by x, then we could add (θ, x) to the correspondence without destroying monotonicity and this would contradict the maximality of ξ. It must also be the case, by maximal monotonicity, that $\xi \neq \emptyset$. Let $(\theta_0, x_0) \in \xi$. Now ϕ_{x_0} supports τ^* at θ_0. If $\tau^*(\theta) - \phi_{x_0}(\theta) \leq \varepsilon$ then we are done, so assume the contrary. By definition, τ^* is ε supported at θ by some elementary function $f(\theta) = v(x', \theta) - t$ for some x', t. We must show that $x' \in \operatorname{im}\xi$, so we can then choose $\theta' \in \xi^{-1}(x')$, $t = \tau(x')$. It cannot be that $x' < x_0$, for if that were so then by single crossing $\tau^*(\theta) - \phi_{x'}(\theta) \geq \tau^*(\theta) - \phi_{x_0}(\theta) > \varepsilon$. So $x' \in [x_0, \infty) \subset \operatorname{im}\xi$. This is the first part of what we need to show. Since $x' \in \operatorname{im}\xi$ there must be some θ' such that $(\theta', x') \in \xi$. So $\phi_{x'}(\theta) = v(x', \theta) - \tau(x')$ supports τ^* at θ'. So $t \geq \tau(x')$, and we can repace t by $\tau(x')$. ■

Proof of Theorem 3. We show first that the conditions imply that τ is supported at all interior points of $\operatorname{dom}\tau$. If $\operatorname{im}\xi \neq \operatorname{dom}\tau$ then there exists $\tilde{x} \in \operatorname{dom}\tau$ such that either $\tilde{x} < \operatorname{im}\xi$ or $\tilde{x} > \operatorname{im}\xi$; we consider the former case.

We show first that $\tau^* - \phi_{\tilde{x}}$ is monotone increasing. Consider $\theta' < \theta'' \in \operatorname{dom}\tau^*$. Choose $(x_0, \theta_0) \in \xi$, and consider the elementary function ϕ_{x_0} that supports τ^* at θ_0. Let $\varepsilon = \left(\phi_{x_0}(\theta'') - \phi_{\tilde{x}}(\theta'')\right) - \left(\phi_{x_0}(\theta') - \phi_{\tilde{x}}(\theta')\right)$ and note that $\varepsilon > 0$ by strict supermodularity, since $\tilde{x} < x_0$. By the Lemma there exists $(x_1, \theta_1) \in \xi$ and an elementary function ϕ_{x_1} that supports τ^* at θ_1 such that $\phi_{x_1}(\theta') \leq \tau^*(\theta') \leq \phi_{x_1}(\theta') + \frac{\varepsilon}{3}$. Without loss of generality we can assume that $\theta_0 \leq \theta_1 \leq \theta'$ (if $\theta' \in \operatorname{dom}\xi$ we can set $\theta_0 = \theta_1 = \theta'$; if $\theta' \notin \operatorname{dom}\xi$ then $\theta_0, \theta_1 < \theta'$ and if $\theta_1 < \theta_0$ we can replace θ_1 by θ_0). Then $\left(\tau^*(\theta'') - \phi_{x_1}(\theta'')\right) - \left(\tau^*(\theta') - \phi_{x_1}(\theta')\right) \geq -\left(\tau^*(\theta') - \phi_{x_1}(\theta')\right) \geq -\frac{\varepsilon}{3}$, and $\left(\phi_{x_1}(\theta'') - \phi_{x_0}(\theta'')\right) - \left(\phi_{x_0}(\theta') - \phi_{x_0}(\theta')\right) \geq 0$ by supermodularity, and $\left(\phi_{x_0}(\theta'') - \phi_{\tilde{x}}(\theta'')\right) - \left(\phi_{x_0}(\theta') - \phi_{\tilde{x}}(\theta')\right) = \varepsilon > 0$, so $\left(\tau^*(\theta'') - \phi_{x_1}(\theta'')\right) - \left(\tau^*(\theta') - \phi_{x_1}(\theta')\right) \geq -\frac{2\varepsilon}{3} > 0$. Thus $\tau^* - \phi_{\tilde{x}}$ is monotone.

We now establish a contradiction. So assume that there exists another such $\tilde{x}'$ with $\tilde{x} < \tilde{x}'$, and consider the elementary function $\phi_{\tilde{x}'}$, which also virtually supports τ^*. It cannot be the case that $\phi_{\tilde{x}'}$ crosses $\phi_{\tilde{x}}$ at any point in $\operatorname{dom}\tau^*$. For assume that it did so, say at θ_1. Then $\phi_{\tilde{x}'}(\theta_1) = \phi_{\tilde{x}}(\theta_1) < \tau^*(\theta_1)$, so $\tau^*(\theta) - \phi_{\tilde{x}'}(\theta)$ is bounded uniformly away from 0 for $\theta \geq \theta_1$ by monotonicity. On the other hand, $\phi_{\tilde{x}'}(\theta_1) < \phi_{\tilde{x}}(\theta_1) < \tau^*(\theta_1)$ for $\theta < \theta_1$ and $\phi_{\tilde{x}}(\theta) - \phi_{\tilde{x}'}(\theta)$ is strictly decreasing in this range. So $\tau^*(\theta) - \phi_{\tilde{x}'}(\theta)$ is bounded uniformly away from 0 on some neighbourhood U of θ_1 by continuity, and by $\phi_{\tilde{x}}(\theta'') - \phi_{\tilde{x}'}(\theta'')$ for $\theta \leq \theta''$, where θ'' is some point in U to the left of θ_1. That is, it is not the case that $\phi'(\theta)$ virtually supports $\tau^*(\theta)$.

So $\phi_{\tilde{x}}$ does not cross $\phi_{\tilde{x}'}$. Assume that $\phi_{\tilde{x}} < \phi_{\tilde{x}'} < \tau^*$. But $\phi_{\tilde{x}}$ virtually supports τ^*, so it must be that $\phi_{\tilde{x}}$ virtually supports $\phi_{\tilde{x}'}$. This contradiction establishes the first part of the Theorem.

We now turn to the converse. So assume that we have two elementary functions $\phi_0(\theta) = v(x_0, \theta) - t_0$ and $\phi_2(\theta) = v(x_2, \theta) - t_2$ such that $\phi_2(\theta) > \phi_0(\theta)$, $\sup_\theta \phi_0(\theta) - \phi_2(\theta) = 0$. Assume that $x_0 < x_2$. This means that the locus of virtual support is at the left of $\operatorname{im}\xi$; the converse case is symmetrical. It also means that the type space Θ cannot have a minimal element, for then the supremum would be a minimum, contradicting the inequality $\phi_2(\theta) > \phi_0(\theta)$. Choose x_1 strictly between x_0 and x_2 and consider $\phi_1(\theta) = v(x_1, \theta) - t_1$, where $t_1 = \inf_\theta (v(x_2, \theta) - v(x_1, \theta)) - t_2$.

It is easy to check that $\phi_1(\theta)$ lies everywhere between $\phi_0(\theta)$ and $\phi_2(\theta)$. Consider the function $\tau(x) = \sup_\theta v(x, \theta) - \phi_2(\theta)$. Then, for $i = 0, 1$, $\tau(x_i) = \sup_\theta v(x_i, \theta) - \phi_2(\theta) = t_i + \sup_\theta \phi_i(\theta) - \phi_2(\theta)$ but the supremum is not attained as a maximum. Thus τ is not supported at x_0 or at x_1. It is clear that we can produce in this way a continuum of points in $\operatorname{dom}\tau$ where τ is not supported. ■

References

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

Berge, C. (1963): *Topological Spaces*. Oliver and Boyd, London.

Bonnans, J., and A. Shapiro (2000): *Perturbation analysis of optimization problems*. Springer Verlag.

Border, K. (1991): "Implementation of Reduced Form Auctions: a Geometric Approach," *Econometrica*, 59, pp. 1175–1187.

——— (2007): "Reduced Form Auctions Revisited," *Economic Theory*, 31, pp. 167–181.

Jehle, G. A., and P. J. Reny (2000): *Advanced Microeconomic Theory*. Addison Wesley, second edn.

Krishna, V., and E. Maenner (2001): "Convex Potentials with an Application to Mechanism Design," *Econometrica*, 69(4), pp. 1113–1119.

Laffont, J., and D. Martimort (2002): *The theory of incentives: the principal-agent model.* Princeton Univ Pr.

Milgrom, P., and I. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70(2), 583–601.

Pallaschke, D., and S. Rolewicz (1997): *Foundations of mathematical optimization*, vol. 388 of *Mathematics and its Applications.* Kluwer Academic Publishers Group, Dordrecht, Convex analysis without linearity.

Rochet, J. C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasilinear Context," *Econometrica*, 16, pp. 191–200.

Rockafellar, R. T. (1970): *Convex analysis*, Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J.

Rockafellar, R. T., and R. J.-B. Wets (1998): *Variational analysis*, vol. 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin.

Rogerson, W. (1987): "On the Optimality of Menus of Linear Contracts," *Northwesten Discussion Paper*, 714R.

Rubinov, A. (2000): *Abstract convexity and global optimization*, vol. 44 of *Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht.

Vohra, R. (2011): *Mechanism Design a Linear Programming Approach.* Cambridge University Press, New Yoprk.