Wirtschaftswissenschaftliches Zentrum (WWZ) der Universität Basel

March 2015

The Implementation Duality

WWZ Discussion Paper 2015/01

Georg Nöldeke, Larry Samuelson

Corresponding author:

Prof. Georg Nöldeke

University of Basel
Faculty of Business and Economics
Peter Merian-Weg 6
CH-4002 Basel

Phone: +41(0)61 267 33 20
georg.noeldeke@unibas.ch

A publication of the Center of Business and Economics (WWZ), University of Basel.

The Implementation Duality*

Georg Nöldeke[†] Larry Samuelson[‡]

March 14, 2015

Abstract: We use the theory of abstract convexity to study adverse-selection principal-agent problems and two-sided matching problems, departing from much of the literature by not requiring quasilinear utility. We formulate and characterize a basic underlying implementation duality. We show how this duality can be used to obtain a sharpening of the taxation principle, to obtain a general existence result for solutions to the principal-agent problem, to show that (just as in the quasilinear case) all increasing decision functions are implementable under a single crossing condition, and to obtain an existence result for stable outcomes featuring positive assortative matching in a matching model.

Keywords: Implementation, Duality, Galois Connection, Imperfectly Transferable Utility, Principal-Agent Model, Two-Sided Matching

***JEL* Classification Numbers:** C62, C78, D82, D86.

*We thank Christian Ewerhart, Samuel Häfner, Bruno Jullien, Igor Letina, Thomas Mariotti, Sofia Moroni, Nick Netzer, Frank Riedel, and Jean-Charles Rochet for helpful comments and discussion. Much of the work of this paper was done while the authors were visiting IAST (Nöldeke) and IDEI (Samuelson) at the University of Toulouse. We are grateful for their hospitality. Larry Samuelson thanks the National Science Foundation (SES-1153893) for financial support.

[†]Faculty of Business and Economics, University of Basel, Switzerland, georg.noeldeke@unibas.ch

[‡]Department of Economics, Yale University, larry.samuelson@yale.edu

The Implementation Duality

Contents

The Implementation Duality

1 Introduction

In this paper we present and pursue a nonlinear-pricing approach to adverse-selection principal-agent problems and two-sided matching problems. The key observation underlying our approach is that there is a natural duality between the utilities different types of agents obtain and the prices attached to different decisions. This is most easily seen in the matching context. Here, utilities received by agents on one side on the market serve a dual role as prices, constraining the options available to agents on the other side of the market. Considering the value functions of the resulting constrained optimization problems defines a pair of functions, mapping utility profiles for agents on one side of the market into utility profiles for agents on the other side of the market. We refer to these functions as implementation maps. The mathematics carry over without any changes to the principal-agent context, where agents are matched with decisions rather than with other agents.

Throughout this paper we work without requiring quasilinear utility functions. In the terminology of matching models we thus allow for imperfectly transferable utility. In such a setting the implementation maps are dualities in the sense of abstract convex analysis (Singer, 1997) or, equivalently, form a Galois connection (Ore, 1944; Birkhoff, 1995). We show that under familiar compactness and continuity assumptions this "implementation duality" has convenient topological properties. Together with the order structure inherent in a Galois connection, these topological properties provide the fuel for our analysis.

Section 2 presents our formal framework and provides the interpretation of our abstract implementation problem both in terms of the adverse-selection principal-agent model and the two-sided matching model. Section 3 introduces the implementation maps and studies the duality associated with these maps, culminating in a general characterization of implementability. In the context of the principal-agent model this characterization strengthens the taxation principle by identifying the set of "implementable tariffs" as a minimal set of nonlinear tariffs required to implement any incentive compatible direct mechanism. Because of the underlying duality structure, the set of implementable utility profiles and implementable tariffs have identical structural properties. We show that both of these sets are closed and that bounded sets of implementable tariffs and utility profiles are uniformly equicontinuous. These results provide the foundations for the approach considered in this paper, and we expect them to find additional uses.

Sections 4–6 develop applications of the results in Section 3, explaining how they fit into the literature as we proceed. Section 4 formulates the principal's optimization problem in the principal-agent model as a nonlinear pricing problem. Due to the possibility of restricting attention to implementable tariffs, the existence of a solution follows from a fairly straightforward application of Weierstrass' extreme value theorem. We also show that in the absence of quasilinearity, the optimal tariff may leave slack in all of the agent's participation constraints, and we identify sufficient conditions (implied by the single-crossing condition of Section 5) for a participation constraint to bind.

Section 5 considers the most commonly studied special case of the principal-agent model in which types and decisions are one-dimensional and the agent's utility function satisfies a single crossing condition. We show that under these conditions every increasing decision function is implementable for any initial condition. Further, if the single crossing condition holds strictly, a decision function is implementable if and only if it is increasing. For the case in which the agent's utility is quasilinear, these results are well-known. Our contribution lies in showing that even in the absence of quasilinearity these results are a direct (albeit nontrivial) consequence of the underlying duality structure. In particular, as in the quasilinear case, no smoothness assumptions on utility functions or decision functions are required.

In the context of a two-sided matching model, the single crossing condition we impose in Section 5 is equivalent to the condition of generalized increasing differences introduced in Legros and Newman (2007a). Section 6 studies a matching model satisfying this condition and shows that the existence of a stable positive assortative matching follows almost immediately from applying our previous results to the dual relationship between the set of implementable utility profiles for the agents on the different sides of the market.

There is a long tradition, going back at least to Rochet (1987), of using duality results from convex analysis (Rockafellar, 1970) and generalized convex analysis in the investigation of principal-agent models and, more generally, in mechanism design.[1] Ekeland (2010a) provides a survey of results from generalized convex analysis that are related to the ones we obtain in Section 3 in the context of a principal agent-model, emphasizing the connection to the theory of optimal transportation (Villani, 2009). Bardsley (2012) offers a different perspective and provides additional references. Vohra (2011) discusses the application of convex analysis to mechanism design; recent notable contributions to this literature include Heydenreich, Müller, Uetz, and Vohra (2009) and Goeree and Kushnir (2013). That there is a common duality structure underlying adverse-selection principal-agent problems and two-sided matching models has been noted most prominently in the literature applying optimal transportation theory to matching problem (cf. Ekeland, 2010b; Chiappori, McCann, and Nesheim, 2010).

Quasilinearity is indispensable for the applicability of the tools from generalized convex analysis and optimal transportation theory. It is thus no coincidence that all of the literature referenced in the preceding paragraph considers the case in which

[1]The cyclical monotonicity condition for incentive compatibility from Rochet (1987) has been refined and extended in many directions. See Carbajal and Ely (2013) for a useful discussion of this literature.

utility functions are quasilinear. While the assumption of quasilinear (or perfectly transferable) utility is standard in the literature on adverse-selection principal-agent models and two-sided matching models, it excludes the possibility of taking income effects into account and, as argued in Legros and Newman (2007a), also precludes matching models in which utility possibility sets are shaped by underlying incentive and enforcement constraints. It is thus important to understand how duality results may be brought to bear in the absence of quasilinearity. In this paper we identify the relevant duality and show how the properties of this duality can be used to extend some important results from the quasilinear to the general case. Of course, working under more general assumptions also has a cost. For instance, in the absence of quasilinearity there is no hope for revenue equivalence results.

As we have noted before, the duality notion from abstract convex analysis that we employ is equivalent to a Galois connection. Singer (1997) is a basic reference for abstract convex analysis, while Erné, Koslowski, Melton, and Strecker (1993) provide a guide to Galois connections. With the exception of Monjardet (1978, 2007) we have not encountered papers studying Galois connections in an economic context. We believe that one reason for this relative neglect of Galois connections in economics is that obtaining substantial implications requires the presence of additional structure, like the compactness and continuity properties that we impose.

2 Model

2.1 Basic Ingredients

The basic ingredients of our model are two sets, X and Y, and a function $\phi : X \times Y \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}}$ denotes the extended real numbers. We offer two interpretations of these basic ingredients.[2]

In the first interpretation X is a set of possible types for an agent, Y is a set of possible decisions to be taken by the agent, and $u = \phi(x, y, v)$ is the utility of an agent of type x, who takes decision y and provides monetary transfer v to a principal. To complete the specification of an adverse-selection *principal-agent model*, it remains to specify a utility function for the principal, her beliefs over the agent's types, and reservation utilities for the agent's types. We do so in Section 4.

In the second interpretation X is a set of buyer types, Y is a set of seller types, and ϕ is the utility frontier describing the feasible payoffs that can be realized in a match between buyer type x and seller type y. That is, $u = \phi(x, y, v)$ is the maximal utility buyer type x can obtain when matched with seller type y and providing utility

[2]Using the extended real numbers in the definition of the function ϕ simplifies the statements of preliminary results in Sections 3.1 and 3.2 and clarifies the relationship of our constructions to the duality literature, but none of our subsequent analysis hinges on this modelling choice. When offering intuitive interpretations, as in the following two paragraphs, we thus write as if $\overline{\mathbb{R}}$ in the definition of ϕ were replaced by $\mathbb{R}$.

v to the seller. Specifying measures on X and Y and reservation utilities for the buyer and seller types, as we do in Section 6, provides the missing ingredients to obtain a two-sided *matching model*.

We maintain the following assumption throughout:

Assumption 1. *The sets X and Y are compact subsets of metric spaces. The function ϕ is continuous, strictly decreasing in its third argument, and satisfies $\phi(x, y, \overline{\mathbb{R}}) = \overline{\mathbb{R}}$ for all $(x, y) \in X \times Y$.*

The assumption that the type space X is compact is standard in principal-agent models, but violated in some applications in finance (such as Glosten, 1989) in which unbounded type spaces are considered. As the latter models are known to behave quite differently (cf. Hellwig, 1992; Mailath and Nöldeke, 2008), we save them for separate consideration. In a matching model, compactness of Y is simply the mirror image of the compactness assumption on X. In the principal-agent model compactness of Y obtains naturally in many situations (e. g. when Y is the set of probability distributions over a finite number of alternatives), but is also applicable in principal-agent problems in which the set of possible decisions is more naturally taken to be unbounded (e. g. the set of possible quantities produced by a monopoly), but conditions on the primitives of the model (e. g. on marginal costs and willingness to pay for the monopolist's product) allow attention to be restricted to a compact subset, which we identify with Y. We view the continuity assumption on ϕ as being uncontroversial, whereas the strict monotonicity of ϕ in its third argument is in line with the interpretation of the third argument as a monetary transfer in the principal-agent model. In the matching model, the strict monotonicity requirement excludes the nontransferable utility case, while allowing for either perfectly or imperfectly transferable utility. Remark 5 in Section 4.3 explains the role and importance of the full range assumption. Section 3.3 places Assumption 1 in the context of the duality literature.

We find it convenient to use a neutral term for ϕ and refer to it (for reasons explained in Section 3.2) as the *generating function*. The conditions on the generating function in Assumption 1 are satisfied if there exists a continuous function $f : X \times Y \to \mathbb{R}$ such that $\phi(x, y, v) = f(x, y) - v$. We refer to this as the *quasilinear case* as it corresponds to assuming that the agent's utility function in a principal-agent model is quasilinear in the transfer. In a matching model, the quasilinear case corresponds to assuming that utility is perfectly transferable between any pair of matched agents.

2.2 The Inverse Generating Function

The *inverse generating function* $\psi : Y \times X \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ is specified as the solution to the equation

$$u = \phi(x, y, \psi(y, x, u)). \tag{1}$$

Assumption 1 ensures that for all $x \in X$, $y \in Y$ and $u \in \overline{\mathbb{R}}$, there is a unique value $v \in \overline{\mathbb{R}}$ satisfying $u = \phi(x, y, v)$, so that the function ψ introduced in (1) is indeed well-defined and satisfies the "reverse" inverse relationship

$$v = \psi(y, x, \phi(x, y, v)). \tag{2}$$

Furthermore, the inverse generating function inherits the properties of the generating function stated in Assumptions 1. The proof in Appendix A.1 shows that the following is a direct consequence of (1):

Lemma 1. *Let Assumption 1 hold. Then ψ is continuous, strictly decreasing in its third argument, and satisfies $\psi(y, x, \overline{\mathbb{R}}) = \overline{\mathbb{R}}$ for all $(y, x) \in Y \times X$.*

In the context of a matching model the interpretation of ψ is analogous to the one given for ϕ: $v = \psi(y, x, u)$ is the maximal utility a seller type y can obtain when matched with a buyer type x and providing utility u to the buyer. Observe that in the definition of ψ the order of the first two arguments has been exchanged, so that in a matching model for both ϕ and ψ the first argument gives the type of the agent whose maximal utility is specified and the second argument gives the type of his or her partner. In particular, in the quasilinear case we have $\psi(y, x, u) = g(y, x) - u$, where $g(y, x) = f(x, y)$ holds for all $(x, y) \in X \times Y$.

In the context of a principal-agent model the function ψ identifies the largest transfer an agent of type x can pay for the decision y while obtaining utility level u.[3] In either context, as indicated by (1)–(2) and Lemma 1, the inverse generating function contains the same information about preferences as the generating function and has identical properties.

2.3 Profiles and Assignments

Let $\overline{\mathbb{R}}^X$ denote the set of functions from X to $\overline{\mathbb{R}}$ and let $\overline{\mathbb{R}}^Y$ be the set of functions from Y to $\overline{\mathbb{R}}$. With the pointwise partial order inherited from the standard order $\geq$ on $\overline{\mathbb{R}}$ the sets $\overline{\mathbb{R}}^X$ and $\overline{\mathbb{R}}^Y$ are complete lattices. For simplicity we denote both of these partial orders by $\geq$. For $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \overline{\mathbb{R}}^Y$ we use the standard notation $\boldsymbol{u}_1 \vee \boldsymbol{u}_2$ and $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$, respectively $\boldsymbol{v}_1 \vee \boldsymbol{v}_2$ and $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$, to denote the joins (obtained by taking pointwise maxima) and meets (obtained by taking pointwise minima).

Throughout the following we refer to $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ as *profiles.* In the context of a matching model we interpret $\boldsymbol{u}$ as a *utility profile* for buyers, whereas $\boldsymbol{v}$

[3] Mirrlees (1986, p. 1231) introduces a counterpart to the inverse generating function in his analysis of the optimal income taxation problem; Hellwig (2010, Proposition 2.6) features an application in the context of a principal-agent model. Luenberger (1992) refers to an analogous (but more general) measure of the agent's willingness-to-pay as the benefit function and establishes duality results linking the benefit function to the expenditure function.

is a utility profile for sellers. In the context of a principal-agent model $\boldsymbol{u}$ is again a utility profile, corresponding to the specification of a rent function for the agent types, whereas we interpret $\boldsymbol{v}$ as a non-linear *tariff* offered by the principal to the agent, with $\boldsymbol{v}(y)$ specifying the transfer to the principal at which any type of agent can purchase decision y.

Allowing for profiles mapping into the extended real numbers is convenient for parts of our analysis because $\overline{\mathbb{R}}^X$ and $\overline{\mathbb{R}}^Y$ are complete lattices. Our main interest, however, is in real-valued profiles that are bounded in the sets $\mathbb{R}^X$ and $\mathbb{R}^Y$. We let $\boldsymbol{B}(X) \subset \mathbb{R}^X$ and $\boldsymbol{B}(Y) \subset \mathbb{R}^Y$ denote the corresponding sets of profiles and for simplicity refer to the elements of these sets as bounded profiles. We endow $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$ with the supremum norm, denoted by $\|\cdot\|$ in both cases, making them complete metric spaces for the induced metric. The sets $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$ are conditionally complete sublattices of $\overline{\mathbb{R}}^X$ resp. $\overline{\mathbb{R}}^Y$.[4] We let $\boldsymbol{CB}(X) \subset \boldsymbol{B}(X)$ and $\boldsymbol{CB}(Y) \subset \boldsymbol{B}(Y)$ denote the sets of continuous and bounded profiles.

Let Y^X denote the set of functions from X to Y and let X^Y be the set of functions from Y to X. Any function $\boldsymbol{y} \in Y^X$ and any function $\boldsymbol{x} \in X^Y$ will be referred to as an *assignment*. In the matching context, the interpretation of an assignment $\boldsymbol{y} \in Y^X$ is that $y = \boldsymbol{y}(x)$ is the type of seller with whom buyer x matches; the interpretation of an assignment $\boldsymbol{x} \in X^Y$ is analogous.[5] In the principal-agent model, an assignment $\boldsymbol{y} \in Y^X$ specifies a decision for every agent type, whereas we can think of an assignment $\boldsymbol{x} \in X^Y$ as specifying for every decision y the agent type $x = \boldsymbol{x}(y)$ to whom the principal assigns this decision. In this context we thus follow Nöldeke and Samuelson (2007) in referring to $\boldsymbol{y}$ as a *decision assignment* and to $\boldsymbol{x}$ as a *type assignment*.

2.4 Implementable Pairs, Profiles, and Assignments

For given $\boldsymbol{u} \in \overline{\mathbb{R}}^X$, respectively for given $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$, define the correspondences $\boldsymbol{X_u} : Y \rightrightarrows X$ and $\boldsymbol{Y_v} : X \rightrightarrows Y$ by

$$\boldsymbol{X_u}(y) = \operatorname*{argmax}_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \tag{3}$$

$$\boldsymbol{Y_v}(x) = \operatorname*{argmax}_{y \in Y} \phi(x, y, \boldsymbol{v}(y)). \tag{4}$$

[4] That is, all nonempty, order bounded subsets of profiles in $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$ have suprema and infima in these sets (cf. Birkhoff, 1995, Chapter V.3).

[5] Note that the definition of an assignment does not incorporate any notion of feasibility (e.g., an assignment $\boldsymbol{x}$ could specify that all types of the seller match with the same type of buyer). In the matching context an assignment is thus sometime referred to as a semi-matching. An appropriate feasibility notion for the matching model will be introduced in Section 6.

We say that a profile $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ *implements* the pair $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ if

$$\boldsymbol{y}(x) \in \boldsymbol{Y}_{\boldsymbol{v}}(x) \tag{5}$$

$$\boldsymbol{u}(x) = \max_{y \in Y} \phi(x, y, \boldsymbol{v}(y)) \tag{6}$$

holds for all $x \in X$ (which, obviously, implies that $\boldsymbol{Y}_{\boldsymbol{v}}(x)$ is non-empty valued). Similarly, a profile $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ implements the pair $(\boldsymbol{v}, \boldsymbol{x}) \in \mathbb{R}^X \times X^Y$ if

$$\boldsymbol{x}(y) \in \boldsymbol{X}_{\boldsymbol{u}}(y) \tag{7}$$

$$\boldsymbol{v}(y) = \max_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \tag{8}$$

holds for all $y \in Y$. We also say that a profile $\boldsymbol{v}$ *implements* the profile $\boldsymbol{u}$ (assignment $\boldsymbol{y}$) if there exists $\boldsymbol{y}$ (there exists $\boldsymbol{u}$) such that $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ and use the analogous terminology for profiles $\boldsymbol{u}$ and assignments $\boldsymbol{x}$.

Note that our definitions require that implementable profiles be real-valued, whereas a profile implementing a pair may take on values in the extended real numbers. Remark 1 at the end of this section explains the reasons for these choices. Later on we show that it is without loss of generality to require both implementing and implemented profiles profiles to be not only real-valued, but also bounded (see Lemma 2).

The interpretation of the above implementation conditions in a matching model is clear: When $\boldsymbol{v}$ implements the pair $(\boldsymbol{u}, \boldsymbol{y})$ every buyer type x finds it optimal to select seller type $\boldsymbol{y}(x)$ as a partner and by doing so obtains the utility $\boldsymbol{u}(x)$, given that sellers have to be provided with the utility profile $\boldsymbol{v}$. Similarly, $\boldsymbol{u}$ implements the pair $(\boldsymbol{v}, \boldsymbol{x})$ if every seller type y finds it optimal to select buyer type $\boldsymbol{x}(y)$ as a partner and obtains utility $\boldsymbol{v}(y)$ from doing so when faced with the utility profile $\boldsymbol{u}$.

The interpretation of the statement that $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ in a principal-agent model is familiar: when faced with the tariff $\boldsymbol{v}$, an agent of type x finds it optimal to choose decision $\boldsymbol{y}(x)$ and obtains utility $\boldsymbol{u}(x)$ from doing so. The interpretation of the statement that $\boldsymbol{u}$ implements $(\boldsymbol{v}, \boldsymbol{x})$ is less standard. It means the following: For every decision $y \in Y$, type $\boldsymbol{x}(y)$ is (one of) the type(s) who can pay the most for the decision y (while receiving utility at least $\boldsymbol{u}(x)$) and $\boldsymbol{v}(y)$ is the corresponding maximized willingness-to-pay for the decision y. Assuming that the principal's utility is strictly increasing in the transfer and (for the sake of illustration) does not depend on the type of agent with whom she interacts, $\boldsymbol{x}(y)$ thus identifies (one of) the most profitable type(s) of the agent to whom the principal would want to sell decision y and $\boldsymbol{v}(y)$ is the resulting revenue, given that type x of the agent has to be provided with utility level $\boldsymbol{u}(x)$.

Pairs, profiles, and assignments are said to be *implementable* if there exists a profile implementing them.

Remark 1 (Implementability and Direct Mechanisms)**.** As noted above, our definitions require that implementable profiles be real-valued, whereas a profile implementing a pair may take on values in the extended real numbers. Together these features ensure that in the principal-agent model a pair $(\boldsymbol{u}, \boldsymbol{y})$ is implementable in the sense defined above if and only if there exists an incentive compatible direct mechanism $(\boldsymbol{t}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$, satisfying $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x))$ for all $x \in X$. The only-if direction of this claim is immediate from the revelation principle, while the if direction follows because standard proofs of the taxation principle (e.g. Guesnerie and Laffont, 1994; Rochet, 1985) are applicable to our setting. In particular, these standard proofs apply when infinite transfers are allowed, motivating us to allow profiles that map into the extended reals. ■

3 Duality

In this section we provide a general characterization of implementable pairs and profiles. Section 3.1 introduces a pair of functions between the sets of profiles that we refer to as implementation maps. These implementation maps are the key building block in all of our subsequent analysis. The implementation maps are dualities (in the sense of Singer, 1997) between the sets of profiles $\overline{\mathbb{R}}^X$ and $\overline{\mathbb{R}}^Y$. Equivalently, the implementation maps give rise to a Galois connection between these sets. Section 3.2 makes these statements precise and provides some further discussion and properties of Galois connections.

Linking the duality structure of the implementation maps to properties of implementable pairs and profiles requires more than the order structure captured by the observation that these maps are a Galois connection. Section 3.3 provides such a link. Here we exploit Assumption 1 to show that there is no loss of generality in restricting the domain of the implementation maps to the set of bounded profiles and that, once this is done, the sets of implementable profiles coincide with the ranges of the implementation maps. Building on these observations, Section 3.4 provides a characterization of implementable profiles and implementable pairs. Proposition 3 shows that every implementable pair can be implemented by an implementable profile. As an immediate consequence, we obtain a strengthening of the taxation principle, asserting that every incentive compatible direct mechanism in the principal-agent model can be obtained as the outcome of facing the agent with an *implementable* tariff (Corollary 3). Section 3.5 illustrates Proposition 3 by relating it to the duality results we have obtained for a quasilinear principal-agent model in Nöldeke and Samuelson (2007). Section 3.6 notes some implications of our characterization results for the topological structure of the set of implementable profiles, culminating in the uniform equicontinuity result (Corollary 5) that we have already noted in Section 1.

3.1 Implementation Maps

We define the *implementation maps* $\Phi : \overline{\mathbb{R}}^Y \to \overline{\mathbb{R}}^X$ and $\Psi : \overline{\mathbb{R}}^X \to \overline{\mathbb{R}}^Y$ by setting

$$\Phi \boldsymbol{v}(x) = \sup_{y \in Y} \phi(x, y, \boldsymbol{v}(y)) \quad \forall x \in X \tag{9}$$

$$\Psi \boldsymbol{u}(y) = \sup_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \quad \forall y \in Y. \tag{10}$$

Observe that $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ implements $\boldsymbol{u} \in \mathbb{R}^X$ if and only if $\boldsymbol{u} = \Phi \boldsymbol{v}$ holds and, in addition, the suprema in (9) are finite and attained for all $x \in X$, that is, $\boldsymbol{Y_v}$ is non-empty valued. Similarly, $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ implements $\boldsymbol{v} \in \mathbb{R}^Y$ if and only if $\boldsymbol{v} = \Psi \boldsymbol{u}$ holds and, in addition, the suprema in (10) are finite and attained for all $y \in Y$, that is, $\boldsymbol{X_u}$ is non-empty valued. Consequently, it is clear that the sets of implementable profiles, which we denote by $\boldsymbol{I}(X) \subset \mathbb{R}^X$ and $\boldsymbol{I}(Y) \subset \mathbb{R}^Y$ in the following, are contained in the images of the implementation maps.

3.2 The Galois Connection

The implementation map Φ is a *duality* in the sense of Singer (1997) with Ψ corresponding to the *dual* of Φ.[6] Penot (2000, 2010) refers to the function ϕ as the generating function for the duality Ψ, explaining our use of the term. Every duality between complete lattices gives rise to a *Galois connection* (Birkhoff, 1995, p. 124) between these lattices (and vice versa, cf. Singer, 1997, Theorem 5.4, p. 179).[7] As the proof is straightforward for generating functions satisfying Assumption 1, we include it here.

Proposition 1. *Let Assumption 1 hold. The implementation maps Φ and Ψ are a Galois connection. That is,*

$$\boldsymbol{u} \ \geq \ \Phi \boldsymbol{v} \iff \boldsymbol{v} \ \geq \ \Psi \boldsymbol{u} \tag{11}$$

holds for all $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$.

[6] Singer (1997, Definition 5.1, p. 172) defines a duality as a map between complete lattices which is order reversing in the sense that for any set (including the empty set), the image of the infimum of that set is the supremum of its image. For the case in which the complete lattices are given by $\overline{\mathbb{R}}^X$ and $\overline{\mathbb{R}}^Y$ it follows from the characterization result in Singer (1997, Theorem 7.4, p. 230) that the implementation map Φ is a duality.

[7] Observe that there is an alternative definition of a Galois connection in which the second inequality in (11) below is reversed (Davey and Priestley, 2002, Chapter 7) and that in the original definition of a Galois connection (Ore, 1944) both inequalities in (12) are reversed. None of these differences are essential, as equivalence between the various definitions can be restored by replacing either $\overline{\mathbb{R}}^Y$ or $\overline{\mathbb{R}}^X$ or both by their order dual(s). The kind of Galois connection we consider is sometimes referred to as an antitone Galois connection.

Proof. To obtain (11) observe:

$$
\begin{aligned}
\boldsymbol{u} \geq \Phi \boldsymbol{v} &\iff \boldsymbol{u}(x) \geq \sup_{y \in Y} \phi(x, y, \boldsymbol{v}(y)) \text{ for all } x \in X \\
&\iff \boldsymbol{u}(x) \geq \phi(x, y, \boldsymbol{v}(y)) \text{ for all } x \in X \text{ and } y \in Y \\
&\iff \psi(y, x, \boldsymbol{u}(x)) \leq \boldsymbol{v}(y) \text{ for all } x \in X \text{ and } y \in Y \\
&\iff \boldsymbol{v}(y) \geq \sup_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \text{ for all } y \in Y \\
&\iff \boldsymbol{v} \geq \Psi \boldsymbol{u},
\end{aligned}
$$

where the first equivalence holds by the definition of $\Phi\boldsymbol{v}$ in (9), the second is from the definition of the supremum, the third uses (2) and that the inverse generating function ψ is strictly decreasing in its third argument (Lemma 1), the fourth is by the definition of the supremum, and the fifth holds by definition of $\Psi\boldsymbol{u}$ in (10). ■

To interpret Proposition 1 consider the matching model. Suppose we have a pair of profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ such that each buyer $x \in X$ is content to obtain $\boldsymbol{u}(x)$ rather than matching with any seller $y \in Y$ and providing that seller with utility $\boldsymbol{v}(y)$, that is, the inequality $\boldsymbol{u} \geq \Phi\boldsymbol{v}$ holds. It is then intuitive that every seller $y \in Y$ would similarly weakly prefer to obtain utility $\boldsymbol{v}(y)$ to matching with any buyer $x \in X$ who insists on receiving utility $\boldsymbol{u}(x)$, that is, the inequality $\boldsymbol{v} \geq \Psi\boldsymbol{u}$ holds. Reversing the roles of buyers and sellers in this argument suggests the equivalence in (11).

The following corollary notes some standard implications of the fact that Φ and Ψ are a Galois connection. Our terms for these implications follow Davey and Priestley (2002, p. 159).[8] Again, we include the straightforward proof.

Corollary 1. *Let Assumption 1 hold. The implementation maps Φ and Ψ*

[1.1] satisfy the cancellation rule, that is, for all $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$:

$$\boldsymbol{v} \geq \Psi\Phi\boldsymbol{v} \text{ and } \boldsymbol{u} \geq \Phi\Psi\boldsymbol{u}; \tag{12}$$

[1.2] are order reversing, that is, for all $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \overline{\mathbb{R}}^Y$:

$$\boldsymbol{v}_1 \geq \boldsymbol{v}_2 \Rightarrow \Phi\boldsymbol{v}_2 \geq \Phi\boldsymbol{v}_1 \text{ and } \boldsymbol{u}_1 \geq \boldsymbol{u}_2 \Rightarrow \Psi\boldsymbol{u}_2 \geq \Psi\boldsymbol{u}_1; \tag{13}$$

[1.3] and satisfy the semi-inverse rule, that is, for all $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$:

$$\Phi\Psi\Phi\boldsymbol{v} = \Phi\boldsymbol{v} \text{ and } \Psi\Phi\Psi\boldsymbol{u} = \Psi\boldsymbol{u}. \tag{14}$$

[8] As noted in Birkhoff (1995, Section 5.8), the properties stated in Corollary 1.1–1.2 below are in fact equivalent to (11) and are sometimes taken to be the definition of a Galois connection (e.g., Singer, 1997, Definition 5.3 and Remark 5.6).

Proof. We use (11) to establish (12)–(14). In each case we prove the first of the two statements; the second statement follows by a dual argument. First, for any $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ we trivially have $\Phi\boldsymbol{v} \geq \Phi\boldsymbol{v}$, so that setting $\boldsymbol{u} = \Phi\boldsymbol{v}$ in (11) yields (12). Second, let $\boldsymbol{v}_1 \geq \boldsymbol{v}_2$. By (12) we have $\boldsymbol{v}_2 \geq \Psi\Phi\boldsymbol{v}_2$ and thus $\boldsymbol{v}_1 \geq \Psi\Phi\boldsymbol{v}_2$. Applying (11) with $\boldsymbol{v} = \boldsymbol{v}_1$ and $\boldsymbol{u} = \Phi\boldsymbol{v}_2$ then gives the consequent of (13). Third, (12) gives $\boldsymbol{v} \geq \Psi\Phi\boldsymbol{v}$. Applying (13) with $\boldsymbol{v}_1 = \boldsymbol{v}$ and $\boldsymbol{v}_2 = \Psi\Phi\boldsymbol{v}$ to this inequality yields $\Phi\Psi\Phi\boldsymbol{v} \geq \Phi\boldsymbol{v}$. To establish (14) it remains to establish the reverse inequality. For every $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ we have $\Psi\Phi\boldsymbol{v} \geq \Psi\Phi\boldsymbol{v}$, so that using $\Psi\Phi\boldsymbol{v}$ in place of $\boldsymbol{v}$ and $\Phi\boldsymbol{v}$ in place of $\boldsymbol{u}$ in (11) yields $\Phi\boldsymbol{v} \geq \Phi\Psi\Phi\boldsymbol{v}$. ■

Among the properties of the implementation maps noted in Corollary 1, the order-reversal property (Corollary 1.2) is the one with the most obvious economic interpretation: in a matching model all agents prefer to choose from a given set of potential partners when the partners have low rather than high reservation utilities. To provide perspective on the other two properties, let us consider the quasilinear case.

In the quasilinear case the definitions of the implementation maps in (9) and (10) reduce to

$$\Phi\boldsymbol{v}(x) = \sup_{y\in Y}\left[f(x,y) - \boldsymbol{v}(y)\right]$$

$$\Psi\boldsymbol{u}(y) = \sup_{x\in X}\left[g(y,x) - \boldsymbol{u}(x)\right],$$

indicating that in this case $\Phi\boldsymbol{v}$ is nothing but the f-conjugate of $\boldsymbol{v}$ and $\Psi\boldsymbol{u}$ is the g-conjugate of $\boldsymbol{u}$ (cf. Ekeland, 2010a, Section 3.2). The properties noted in Corollary 1 may then be viewed as generalizing corresponding properties from the theory of conjugate duality. Indeed, the cancellation property (Corollary 1.1) corresponds to the statement that the biconjugate of any function is smaller than the function itself and the semi-inverse rule (Corollary 1.3) corresponds to the statement that a conjugate function is its own biconjugate. These are well-known implications of conjugate duality (cf. Ekeland, 2010a, Section 3.4).

The properties of Galois connections stated in Corollary 1 are the ones which are of direct relevance for our subsequent analysis. Galois connections have many other interesting structural properties which are likely to be of relevance in the study of implementation problems. The following remark notes one of these which, again, might be viewed as a generalization of a well-known result from the theory of conjugate duality.

Remark 2 (Fenchel Inequalities)**.** Consider profiles $\boldsymbol{u} \in \overline{\mathbb{R}}^X$ and $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ satisfying the equality $\boldsymbol{u} = \Phi\boldsymbol{v}$. The proof of Proposition 1 shows that we then have

$$\boldsymbol{v}(y) \geq \psi(y, x, \boldsymbol{u}(x)) = \psi(y, x, \Phi\boldsymbol{v}(x)) \tag{15}$$

for all $(x, y) \in X \times Y$. Dually, the equality $\boldsymbol{v} = \Psi\boldsymbol{u}$ implies

$$\boldsymbol{u}(x) \geq \phi(x, y, \boldsymbol{v}(y)) = \phi(x, y, \Psi\boldsymbol{u}(y)) \tag{16}$$

for all $(x, y) \in X \times Y$. In the terminology of Singer (1997, Definition 7.2, p. 236) these are the generalized Fenchel-Young inequalities. Dropping the middle term in each of (15) and (16), we can rewrite the remaining inequalities in the quasilinear case as

$$\begin{aligned} \boldsymbol{v}(y) + \Phi \boldsymbol{v}(x) &\geq f(x, y) \\ \boldsymbol{u}(x) + \Psi \boldsymbol{u}(y) &\geq g(y, x), \end{aligned}$$

for all $(x, y) \in X \times Y$. These are familiar Fenchel inequalities for conjugate functions (see, for instance, Ekeland, 2010a, Proposition 7), which play an important role in convex analysis because the graphs of the argmax-correspondences $\boldsymbol{X_u}$ and $\boldsymbol{Y_v}$ coincide with the subsets of $X \times Y$ for which the corresponding Fenchel inequality holds with equality. This observation carries over to the generalized Fenchel-Young inequalities given above. ■

3.3 Linking Implementable Profiles and the Implementation Maps

Singer (1997, Theorem 7.4, p. 230) shows that Proposition 1 and, thus, Corollary 1 hold under assumptions on the generating function weaker than the ones we have imposed in Assumption 1. In particular, the sets X and Y may be arbitrary, the only continuity property required is lower semicontinuity of the generating function in its third argument, and weak rather than strict monotonicity in the third argument suffices. The role of Assumption 1 in our analysis thus is not to ensure that the implementation maps are a Galois connection. Rather, it ensures that we have a particularly conveniently-behaved duality.

We begin by noting that due to Assumption 1 every implementable pair features a bounded profile and can in turn be implemented by a bounded profile. Intuitively, one cannot implement a profile $\boldsymbol{u}$ of buyer utilities in which two buyers receive utilities that are arbitrarily far apart because the low-utility buyer would then prefer the seller chosen by the high-utility buyer. Appendix A.2 proves:

Lemma 2. *Let Assumption 1 hold. If* $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ *is implementable, then* $\boldsymbol{u} \in \boldsymbol{B}(X)$ *holds and there exists* $\boldsymbol{v} \in \boldsymbol{B}(Y)$ *implementing* $(\boldsymbol{u}, \boldsymbol{y})$*. Dually, if* $(\boldsymbol{v}, \boldsymbol{x}) \in \mathbb{R}^Y \times X^Y$ *is implementable, then* $\boldsymbol{v} \in \boldsymbol{B}(Y)$ *holds and there exists* $\boldsymbol{u} \in \boldsymbol{B}(X)$ *implementing* $(\boldsymbol{v}, \boldsymbol{x})$*.*

For the purpose of characterizing implementable profiles and pairs we may thus restrict attention to bounded profiles both in the domain and in the range of the implementation maps. From hereon we accordingly take the domains and ranges of the implementation maps Φ and Ψ to be the sets of bounded profiles. With a slight abuse of notation, we continue to denote these maps by Φ and Ψ. Similarly, we will hereafter view sets of profiles as subsets of the metric spaces $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$, so that (for example) we will establish closure properties relative to these sets.

Our next result, proven in Appendix A.3, provides a simple sufficient condition under which a bounded profile implements its image under the relevant implementation map, that is, the suprema in the definitions of the implementation maps are finite and attained. Intuitively, the lower semicontinuity of $\boldsymbol{v}$ ensures that $\phi(x, y, \boldsymbol{v}(y))$ is an upper semicontinuous function of y for each x, at which point Weierstrass' extreme value theorem ensures that it has a maximizer.

Lemma 3. *Let Assumption 1 hold. If a profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is lower semicontinuous, then it implements $\Phi\boldsymbol{v}$. Dually, if a profile $\boldsymbol{u} \in \boldsymbol{B}(X)$ is lower semicontinuous, then it implements $\Psi\boldsymbol{u}$.*

Bounded profiles that are not lower semicontinuous may fail to implement their image under the relevant implementation map. However, replacing a profile by its lower semicontinuous hull (which is bounded) leaves the image of the profile unchanged and, by Lemma 3, provides a profile implementing the given image.[9] In conjunction with Lemma 2 this argument yields the first part of the following lemma. Having established that every implementable profile can be implemented by a profile which is both bounded and lower semicontinuous, the second part of the lemma is a consequence of Berge's maximum theorem. The proof for both parts is in Appendix A.4.

Lemma 4. *Let Assumption 1 hold.*

[4.1] A profile is implementable if and only if it is the image of a bounded profile under the relevant implementation map. That is,

$$\boldsymbol{I}(X) = \Phi\boldsymbol{B}(Y) \text{ and } \boldsymbol{I}(Y) = \Psi\boldsymbol{B}(X). \tag{17}$$

[4.2] All implementable profiles are continuous. That is,

$$\boldsymbol{I}(X) \subset \boldsymbol{CB}(X) \text{ and } \boldsymbol{I}(Y) \subset \boldsymbol{CB}(Y). \tag{18}$$

3.4 Characterizing Implementable Profiles and Pairs

Substituting the condition $\boldsymbol{u} = \Phi\boldsymbol{v}$ into the first statement of the semi-inverse rule in Corollary 1.3 shows that an implementable profile $\boldsymbol{u} \in \boldsymbol{I}(X)$ is not only bounded (Lemma 2) but also satisfies $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$. Vice versa, a profile $\boldsymbol{u} \in \boldsymbol{B}(X)$ satisfying $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$ is the image of the profile $\boldsymbol{v} = \Psi\boldsymbol{u}$ under the implementation map Φ. Because the profile $\boldsymbol{v}$ is bounded and continuous (Lemma 4) it then implements $\boldsymbol{u}$ (Lemma 3), showing that $\boldsymbol{u}$ is implementable. Dually, a profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is implementable if and only if it is bounded and satisfies $\boldsymbol{v} = \Psi\Phi\boldsymbol{v}$. Hence, it is a direct implication of Corollary 1 and the properties of implementable profiles established in

[9]The lower semicontinuous hull of a function is also known as its lsc regularization or its lower closure (Rockafellar and Wets, 1998, p. 14). Our terminology follows Penot (2013, Proposition 1.21).

the previous subsection that implementable profiles are those bounded profiles for which equality holds in the cancellation rule:[10]

Proposition 2. *Let Assumption 1 hold.*

[2.1] Profile $\boldsymbol{u} \in \boldsymbol{B}(X)$ is implementable if and only if $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$.

[2.2] Profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is implementable if and only if $\boldsymbol{v} = \Psi\Phi\boldsymbol{v}$.

We note that for every Galois connection the fixed point condition $u = \Phi\Psi u$ characterizes the image of the map Φ and the fixed point condition $v = \Psi\Phi v$ characterizes the image of the map Ψ (cf. Singer, 1997, Corollary 5.6). In the absence of Assumption 1 there is no guarantee, however, that profiles satisfying these fixed point conditions, even when bounded, are implementable, so that Assumption 1 plays an essential role in going from a characterization of the images of the implementation maps (which we are not interested in as such) to the characterization of implementable profiles in Proposition 2.

It is an immediate implication of Proposition 2 that the restrictions of the implementation maps to the domains $\boldsymbol{I}(X)$ and $\boldsymbol{I}(Y)$ of implementable profiles are inverse to each other and thus are bijections:

$$\boldsymbol{u} = \Phi\boldsymbol{v} \iff \boldsymbol{v} = \Psi\boldsymbol{u}, \quad \text{for all } \boldsymbol{u} \in \boldsymbol{I}(X) \text{ and } \boldsymbol{v} \in \boldsymbol{I}(Y). \tag{19}$$

Using Lemmas 3 and 4.2, every implementable profile $\boldsymbol{u}$ is thus implemented by the implementable profile $\boldsymbol{v} = \Psi\boldsymbol{u}$ and, dually, every implementable profile $\boldsymbol{v}$ is implemented by the implementable profile $\boldsymbol{u} = \Phi\boldsymbol{v}$.[11] Figure 1 provides an illustration.

For the purpose of studying implementable profiles, there is thus nothing lost in eliminating non-implementable profiles from the domain of the implementation maps. The following characterization result shows that the same applies to implementable *pairs*.

Proposition 3. *Let Assumption 1 hold.*

[3.1] If $(\boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{B}(X) \times Y^X$ is implementable, then there is a unique implementable profile $\boldsymbol{v}$ implementing it, namely $\boldsymbol{v} = \Psi\boldsymbol{u}$.

[3.2] If $(\boldsymbol{v}, \boldsymbol{x}) \in \boldsymbol{B}(Y) \times X^Y$ is implementable, then there is a unique implementable profile $\boldsymbol{u}$ implementing it, namely $\boldsymbol{u} = \Phi\boldsymbol{v}$.

Proof. We consider the first claim; the proof for the second is analogous. Let $(\boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{B}(X) \times Y^X$ be implementable. Let $\boldsymbol{v} = \Psi\boldsymbol{u}$. Then, as we have already mentioned, $\boldsymbol{v}$ implements $\boldsymbol{u}$: As the image of a bounded profile under the implementation map Ψ, $\boldsymbol{v}$ is implementable, bounded, and continuous (Lemma 4). From

[10]In the quasilinear case this is the statement that a profile is implementable if and only if it is bounded and is its own generalized biconjugate.

[11]In words we might rephrase this observation as "for every implementable profile there is an implementable profile implementing it, namely the profile implemented by it." This may seem like a mouthful, but such is the nature of a Galois connection.

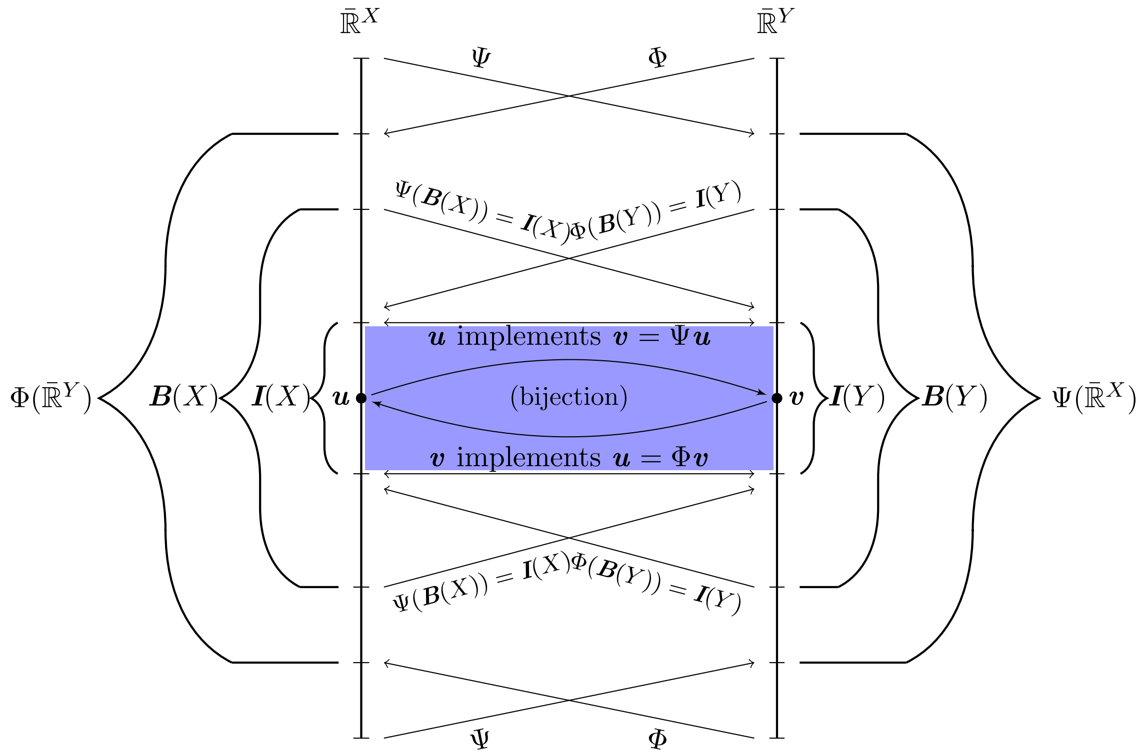

Figure 1: Illustration of the implementation maps. The implementation map Φ maps $\overline{\mathbb{R}}^Y$ into $\overline{\mathbb{R}}^X$ (and Ψ maps from $\overline{\mathbb{R}}^X$ into $\overline{\mathbb{R}}^Y$). When restricted to the set of bounded profiles $B(Y)$, Φ maps into the set of implementable profiles ($I(X)$ (and Ψ maps from the set of bounded profiles $B(X)$ into the set of implementable profiles $I(Y)$). The maps Φ and Ψ are (inverse) bijections on the sets of implementable profiles $I(X)$ and ($I(Y)$ with profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ in these sets satisfying $\boldsymbol{u} = \Phi\boldsymbol{v} \Longleftrightarrow \boldsymbol{v} = \Psi\boldsymbol{u}$ implementing each other.

Lemma 3, $\boldsymbol{v}$ thus implements $\Phi\boldsymbol{v}$. From Proposition 2 we have $\Phi\boldsymbol{v} = \boldsymbol{u}$, so that $\boldsymbol{v}$ implements $\boldsymbol{u}$. Further, from (19) there is no implementable $\tilde{\boldsymbol{v}} \neq \boldsymbol{v}$ implementing $\boldsymbol{u}$. It thus remains to show that $\boldsymbol{v}$ also implements $\boldsymbol{y}$.

As $(\boldsymbol{u}, \boldsymbol{y})$ is implementable there exists $\tilde{\boldsymbol{v}} \in \boldsymbol{B}(Y)$ implementing it, thus satisfying $\boldsymbol{u} = \Phi\tilde{\boldsymbol{v}}$. From the first inequality in the cancellation rule (12) in Corollary 1.1, we have $\tilde{\boldsymbol{v}} \geq \Psi\boldsymbol{u}$ and, thus, $\tilde{\boldsymbol{v}} \geq \boldsymbol{v}$. Now suppose that $\boldsymbol{v}$ does not implement $\boldsymbol{y}$. Because $\boldsymbol{v}$ implements $\boldsymbol{u}$ there then exists $(\hat{x}, \hat{y}) \in X \times Y$ such that

$$\boldsymbol{u}(\hat{x}) = \phi(\hat{x}, \hat{y}, \boldsymbol{v}(\hat{y})) > \phi(\hat{x}, \boldsymbol{y}(\hat{x}), \boldsymbol{v}(\boldsymbol{y}(\hat{x}))) \geq \phi(\hat{x}, \boldsymbol{y}(\hat{x}), \tilde{\boldsymbol{v}}(\boldsymbol{y}(\hat{x}))),$$

where the last inequality uses $\tilde{\boldsymbol{v}} \geq \boldsymbol{v}$ and the assumption that ϕ is decreasing in its third argument. But because $\tilde{\boldsymbol{v}}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ we also have

$$\boldsymbol{u}(\hat{x}) = \phi(\hat{x}, \boldsymbol{y}(\hat{x})), \tilde{\boldsymbol{v}}(\boldsymbol{y}(\hat{x}))),$$

resulting in a contradiction which finishes the proof. ■

It is trivial that $(\boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{B}(X) \times Y^X$ is implementable if it is implemented by $\boldsymbol{v} = \Psi \boldsymbol{u}$, so that Proposition 3.1 does indeed provide a complete characterization of implementable pairs $(\boldsymbol{u}, \boldsymbol{y})$. Similarly, Proposition 3.2 characterizes implementable $(\boldsymbol{v}, \boldsymbol{x}) \in \boldsymbol{B}(Y) \times X^Y$ as those pairs $(\boldsymbol{v}, \boldsymbol{x})$ which are implemented by $\boldsymbol{u} = \Phi \boldsymbol{v}$. We may thus view Proposition 3 as providing a (conceptually) simple test for deciding whether a pair is implementable: given any pair, determine the profile implemented by the profile in that pair. If that profile exists and implements the pair, then the pair is implementable; otherwise it is not.

As we have already established in Lemma 4.2 that implementable profiles are continuous, the following is an immediate consequence of Proposition 3.[12]

Corollary 2. *Every implementable pair can be implemented with a bounded and continuous profile.*

In Remark 1 we have noted that if $(\boldsymbol{t}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ is an incentive compatible direct mechanism, then the pair $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ is implementable, where $\boldsymbol{u}$ is the utility profile resulting when all agent types take decision $\boldsymbol{y}(x)$ and pay the transfer $\boldsymbol{t}(x)$, that is, $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x))$ holds for all $x \in X$. If the tariff $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$, we also have $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$, so that $\boldsymbol{t}(x) = \boldsymbol{v}(\boldsymbol{y}(x))$ holds for all $x \in X$. In conjunction with Proposition 3 we thus have the following strengthening of the taxation principle, asserting that every incentive compatible direct mechanism can be obtained as the the result of facing the agent with an *implementable* tariff $\boldsymbol{v}$:

Corollary 3 (A Stronger Taxation Principle)**.** *Let Assumption 1 hold. If* $(\boldsymbol{t}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ *is an incentive compatible direct mechanism, then there exists a tariff* $\boldsymbol{v} \in \boldsymbol{I}(Y)$ *implementing* $\boldsymbol{y}$ *and satisfying* $\boldsymbol{t}(x) = \boldsymbol{v}(\boldsymbol{y}(x))$ *for all* $x \in X$*.*

Remark 3 (Implementable Behavior)**.** If $\boldsymbol{v} \in \boldsymbol{I}(Y)$ implements $\boldsymbol{u} \in \boldsymbol{I}(X)$, then it follows from Proposition 3 and the equivalence in (19) that $\boldsymbol{u}$ also implements $\boldsymbol{v}$ (and vice versa). This inverse relationship extends to the argmax-correspondences associated with these two profiles: If $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other, then the correspondences $\boldsymbol{X}_{\boldsymbol{u}}$ and $\boldsymbol{Y}_{\boldsymbol{v}}$ satisfy

$$\hat{x} \in \boldsymbol{X}_{\boldsymbol{u}}(\hat{y}) \iff \hat{y} \in \boldsymbol{Y}_{\boldsymbol{v}}(\hat{x}). \tag{20}$$

[12] Weibull (1989) has obtained related results in an optimal taxation model with one-dimensional types and decisions.

To prove (20), we may proceed as follows: As $\boldsymbol{u}$ implements $\boldsymbol{v}$, we have $\boldsymbol{v}(y) = \max_{x \in X} \psi(y, x, \boldsymbol{u}(x))$ for all $y \in Y$, so that the equivalence

$$\hat{x} \in \boldsymbol{X}_{\boldsymbol{u}}(\hat{y}) \iff \boldsymbol{v}(\hat{y}) = \psi(\hat{y}, \hat{x}, \boldsymbol{u}(\hat{x}))$$

obtains. Similarly, as $\boldsymbol{v}$ implements $\boldsymbol{u}$ we have $\boldsymbol{u}(x) = \max_{y \in Y} \phi(x, y, \boldsymbol{v}(y))$ for all $x \in X$ and thus the equivalence

$$\hat{y} \in \boldsymbol{Y}_{\boldsymbol{v}}(\hat{x}) \iff \boldsymbol{u}(\hat{x}) = \phi(\hat{x}, \hat{y}, \boldsymbol{v}(\hat{y})).$$

From the inverse relationship between ϕ and ψ we have

$$\boldsymbol{v}(\hat{y}) = \psi(\hat{y}, \hat{x}, \boldsymbol{u}(\hat{x})) \iff \boldsymbol{u}(\hat{x}) = \phi(\hat{x}, \hat{y}, \boldsymbol{v}(\hat{y})).$$

Combining these three equivalences, we obtain (20).

When $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other, the correspondences $\boldsymbol{X}_{\boldsymbol{u}}$ and $\boldsymbol{Y}_{\boldsymbol{v}}$ are both nonempty-valued. Hence, (20) implies that both of these correspondences are in fact onto, that is, $\boldsymbol{X}_{\boldsymbol{u}}(Y) = X$ and $\boldsymbol{Y}_{\boldsymbol{v}}(X) = Y$ holds. ■

3.5 Example

To illustrate Proposition 3 we consider the following example, for which Nöldeke and Samuelson (2007) have obtained related characterization results. Let $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$ be compact intervals in the real numbers and let the generating function be given by $\phi(x, y, v) = f(x, y) - v$, where the function $f : X \times Y$ is twice continuously differentiable with partial derivatives f_x and f_y and strictly positive cross derivative f_{xy}, that is $f_{xy}(x, y) > 0$ holds for all (x, y) in the interior of $X \times Y$. In the principal-agent interpretation we are thus considering the case in which the utility function of the agent is quasilinear, satisfies standard differentiability conditions, and is strictly supermodular.

In this case it is well-known that a pair $(\boldsymbol{u}, \boldsymbol{y})$ is implementable if and only if $\boldsymbol{y}$ is increasing (Rochet, 1987) and $\boldsymbol{u}$ satisfies (Milgrom and Segal, 2002, Footnote 10)

$$\boldsymbol{u}(x_1) - \boldsymbol{u}(\underline{x}) = \int_{\underline{x}}^{x_1} f_x(x, \boldsymbol{y}(x))dx, \quad \forall x_1 \in X. \tag{21}$$

The set $\boldsymbol{I}(X)$ is then given by the set of profiles $\boldsymbol{u}$ satisfying (21) for some increasing decision assignment $\boldsymbol{y}$.

By symmetry, a pair $(\boldsymbol{v}, \boldsymbol{x})$ is implementable if and only if $\boldsymbol{x}$ is increasing and $\boldsymbol{v}$ satisfies

$$\boldsymbol{v}(y_1) - \boldsymbol{v}(\underline{y}) = \int_{\underline{y}}^{y_1} g_y(y, \boldsymbol{x}(y))dy, \quad \forall y_1 \in Y \tag{22}$$

and the set $\boldsymbol{I}(Y)$ is then given by the set of profiles $\boldsymbol{v}$ satisfying (22) for some increasing type assignment $\boldsymbol{x}$.[13]

For this example, Lemma 1 in Nöldeke and Samuelson (2007) establishes that every implementable $(\boldsymbol{u},\boldsymbol{y})$ can be implemented by a tariff $\boldsymbol{v}$ satisfying (22) for some increasing $\boldsymbol{x}$ and thus, in the terminology we use here, by an implementable tariff. Proposition 3.1 provides a general version of this result and adds the observation that the unique implementable tariff implementing $(\boldsymbol{u},\boldsymbol{y})$ is given by $\boldsymbol{v}=\Psi\boldsymbol{u}$. Using the equivalence $\boldsymbol{v}=\Psi\boldsymbol{u} \Longleftrightarrow \boldsymbol{u}=\Phi\boldsymbol{v}$ in (19) and Lemma 2 in Nöldeke and Samuelson (2007) we can give an explicit expression for this tariff $\boldsymbol{v}$. In particular, $\boldsymbol{v}$ is given by (22), where $\boldsymbol{x}$ is any inverse of $\boldsymbol{y}$ and $\boldsymbol{v}(\underline{y})$ is determined by the initial condition $\boldsymbol{u}(\underline{x})+\boldsymbol{v}(\underline{y})=f(\underline{x},\underline{y})$.[14] Dually, Proposition 3.2 here is the statement that every implementable $(\boldsymbol{v},\boldsymbol{x})$ is implemented by $\boldsymbol{u}$ satisfying (21) with $\boldsymbol{y}$ being an inverse of $\boldsymbol{x}$ and $\boldsymbol{u}(\underline{x})=f(\underline{x},\underline{y})-\boldsymbol{v}(\underline{y})$. Further, there is no other implementable tariff implementing $(\boldsymbol{v},\boldsymbol{x})$.

Nöldeke and Samuelson (2007) point to one use for these results. The standard technique of solving for $(\boldsymbol{u},\boldsymbol{y})$ by examining a relaxed problem that omits the monotonicity constraint on $\boldsymbol{y}$ may yield a solution that violates the omitted constraint, forcing one into an ironing procedure. As we note in Nöldeke and Samuelson (2007), ironing may be avoided by instead examining $(\boldsymbol{v},\boldsymbol{x})$.

[13]Because $f(x,y)=g(y,x)$ holds, we can rewrite (22) as

$$\boldsymbol{v}(y_1)-\boldsymbol{v}(\underline{y})=\int_{\underline{y}}^{y_1} f_y(\boldsymbol{x}(y),y)dy,\quad \forall y_1\in Y,$$

which is the counterpart to equation (5) in Nöldeke and Samuelson (2007).

[14]Given an increasing assignment $\boldsymbol{y}:X\to Y$ let $\boldsymbol{Y}:X\rightrightarrows Y$ denote the maximally monotone correspondence generated by $\boldsymbol{y}$, that is (c.f Rockafellar and Wets, 1998, Exercise 12.9.b),

$$\boldsymbol{Y}(x)=\left[\lim_{x_n\uparrow x}\boldsymbol{y}(x_n),\lim_{x_n\downarrow x}\boldsymbol{y}(x_n)\right]$$

for all $x\in X$ and define the correspondence $\boldsymbol{X}:Y\rightrightarrows X$ for increasing $\boldsymbol{x}:Y\to X$ in an analogous manner. We say that increasing assignments $\boldsymbol{x}$ and $\boldsymbol{y}$ are inverse if and only if the associated correspondences $\boldsymbol{X}$ and $\boldsymbol{Y}$ are inverse, that is, $x\in\boldsymbol{X}(y)\Longleftrightarrow y\in\boldsymbol{Y}(x)$ holds. Let $\boldsymbol{v}$ satisfy (22) for increasing $\boldsymbol{x}$. Now consider any implementable $(\boldsymbol{u},\boldsymbol{y})$, take an inverse $\boldsymbol{x}$ of $\boldsymbol{y}$, and let $\boldsymbol{v}$ satisfy (22) with $\boldsymbol{v}(\underline{y})=f(\underline{x},\underline{y})-\boldsymbol{u}(\underline{x})$. Lemma 2 in Nöldeke and Samuelson (2007) shows that then $\boldsymbol{Y_v}$ is the inverse of $\boldsymbol{X}$, so that $\boldsymbol{v}$ implements $\boldsymbol{x}$ and $\boldsymbol{Y_v}$ is maximally monotone. Further, because $\boldsymbol{Y_v}$ is maximally monotone, we have $\underline{y}\in\boldsymbol{Y_v}(\underline{x})$, implying that the profile $\boldsymbol{u}'$ implemented by $\boldsymbol{y}$ satisfies (21) with $\boldsymbol{u}'(\underline{x})=f(\underline{x},\underline{y})-\boldsymbol{v}(\underline{y})=\boldsymbol{u}(\underline{x})$ and is thus equal to $\boldsymbol{u}$. More generally, under the assumptions maintained in this subsection the profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other if and only if (i) $\boldsymbol{u}(\underline{x})+\boldsymbol{v}(\underline{y})=f(\underline{x},\underline{y})$ or, equivalently, $\boldsymbol{u}(\underline{x})+\boldsymbol{v}(\underline{y})=g(\underline{y},\underline{x})$ holds and (ii) conditions (21) and (22) hold for increasing $\boldsymbol{y}$ and $\boldsymbol{x}$ that are inverse to each other. In this case we have $\boldsymbol{X}=\boldsymbol{X_u}$ and $\boldsymbol{Y}=\boldsymbol{Y_v}$, so that $\boldsymbol{X_u}$ and $\boldsymbol{Y_v}$ are inverse to each other. Remark 3 shows that the latter property does not require the special structure of the example, but holds in general.

3.6 Properties of Implementable Profiles

Proposition 3 provides a powerful tool to obtain properties of (the sets of) implementable pairs and profiles. Here we present two corollaries, which illustrate this for the case of implementable profiles and are used repeatedly in the following. Subsequent sections return to the consideration of implementable pairs.

Besides Proposition 3 we require two topological properties of the implementation maps. These are stated in the following lemma. Appendix A.5 proves:[15]

Lemma 5. *Let Assumption 1 hold. The implementation maps* $\Phi : \boldsymbol{B}(Y) \to \boldsymbol{B}(X)$ *and* $\Psi : \boldsymbol{B}(X) \to \boldsymbol{B}(Y)$ *are*

[5.1] continuous

[5.2] and map bounded sets into bounded sets.

Using Lemma 5.1, we first establish that the sets of implementable profiles are closed in the set of bounded profiles.

Corollary 4. *Let Assumption 1 hold. The sets of implementable profiles* $\boldsymbol{I}(X)$ *and* $\boldsymbol{I}(Y)$ *are closed subsets of* $\boldsymbol{B}(X)$ *resp.* $\boldsymbol{B}(Y)$*.*

Proof. We show that $\boldsymbol{I}(X) \subset \boldsymbol{B}(X)$ is closed; the argument for the dual case is analogous. Consider a sequence $\{\boldsymbol{u}_n\}_{n=1}^{\infty}$ of profiles in $\boldsymbol{I}(X)$ converging to some $\boldsymbol{u}^* \in \boldsymbol{B}(X)$ (in the sup norm). We want to show that $\boldsymbol{u}^*$ is implementable. For all $n \in \mathbb{N}$ we have $\boldsymbol{v}_n = \Psi \boldsymbol{u}_n \in \boldsymbol{I}(Y) \subset \boldsymbol{B}(Y)$, so that $\{\boldsymbol{v}_n\}_{n=1}^{\infty}$ is a sequence in $\boldsymbol{B}(Y)$. By the continuity property of Ψ established in Lemma 5.1, the sequence $\{\boldsymbol{v}_n\}_{n=1}^{\infty}$ converges to $\boldsymbol{v}^* = \Psi \boldsymbol{u}^*$, which is implementable (Lemma 4.1). Because $\boldsymbol{v}_n$ implements $\boldsymbol{u}_n$ (Proposition 3) we have $\boldsymbol{u}_n = \Phi \boldsymbol{v}_n$ for all n. Continuity of Φ then implies $\boldsymbol{u}^* = \Phi \boldsymbol{v}^*$, establishing that $\boldsymbol{u}^*$ is implementable (Lemma 4.1). ■

Next, implementable profiles are not only continuous (as we have already established in Lemma 4.2), but bounded sets of implementable profiles are uniformly equicontinuous. The proof uses Lemma 5.2.

Corollary 5. *Let Assumption 1 hold. If* $\mathcal{U} \subset \boldsymbol{I}(X)$ *is bounded, then it is uniformly equicontinuous. Dually, if* $\mathcal{V} \subset \boldsymbol{I}(Y)$ *is bounded, then it is uniformly equicontinuous.*

[15] It is easy to see that $\mathcal{U} \subset \boldsymbol{B}(X)$ is order bounded (that is, there exists $\underline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ and $\overline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ such that $\underline{\boldsymbol{u}} \leq \boldsymbol{u} \leq \overline{\boldsymbol{u}}$ holds for all $\boldsymbol{u} \in \mathcal{U}$) if and only if it is norm bounded in $\boldsymbol{B}(X)$ (that is, there exists $u \in \mathbb{R}$ such that $\|\boldsymbol{u}\| \leq u$ holds for all $\boldsymbol{u} \in \mathcal{U}$) and that the same observation applies to $\mathcal{V} \subset \boldsymbol{B}(Y)$. Hence, here and throughout the following we simply refer to a set of profiles in $\boldsymbol{B}(X)$ or $\boldsymbol{B}(Y)$ as being bounded without distinguishing between boundedness in order and boundedness in norm.

Proof. Let $\mathcal{V} \subset \boldsymbol{I}(Y)$ be bounded. (The proof for the dual case in which $\mathcal{U} \subset \boldsymbol{I}(X)$ is bounded is analogous.) Fix $\epsilon > 0$. To show uniform equicontinuity of $\mathcal{V}$, we have to establish that there exists $\delta > 0$ such that

$$\|\tilde{y} - y\| < \delta \Longrightarrow \|\boldsymbol{v}(\tilde{y}) - \boldsymbol{v}(y)\| < \epsilon \tag{23}$$

holds for all $\tilde{y}, y \in Y$ and $\boldsymbol{v} \in \mathcal{V}$.

Because $\mathcal{V}$ is bounded, so is $\mathcal{U} = \Phi\mathcal{V}$ (Lemma 5.2). We may then choose $\underline{u} < \bar{u} \in \mathbb{R}$ such that $\boldsymbol{u} \in \mathcal{U}$ implies $\underline{u} \leq \boldsymbol{u}(x) \leq \bar{u}$ for all $x \in X$. Because ψ is continuous (Lemma 1), it is uniformly continuous on the compact set $Y \times X \times [\underline{u}, \bar{u}]$. Consequently, there exists $\delta > 0$ such that

$$\|\tilde{y} - y\| < \delta \Longrightarrow \|\psi(\tilde{y}, x, u) - \psi(y, x, u)\| < \epsilon \tag{24}$$

holds for all $(x, u) \in X \times [\underline{u}, \bar{u}]$. Fix such a δ and let $\|\tilde{y} - y\| < \delta$ hold.

Consider any $\boldsymbol{v} \in \mathcal{V}$. From Proposition 3 the profile $\boldsymbol{u} = \Phi\boldsymbol{v} \in \mathcal{U}$ implements $\boldsymbol{v}$. Let $x \in \boldsymbol{X}_{\boldsymbol{u}}(y)$ and $\tilde{x} \in \boldsymbol{X}_{\boldsymbol{u}}(\tilde{y})$. We then have

$$\begin{aligned}
\boldsymbol{v}(y) =& \psi(y, x, \boldsymbol{u}(x)) \geq \psi(y, \tilde{x}, \boldsymbol{u}(\tilde{x})), \\
\boldsymbol{v}(\tilde{y}) =& \psi(\tilde{y}, \tilde{x}, \boldsymbol{u}(\tilde{x})) \geq \psi(\tilde{y}, x, \boldsymbol{u}(x)),
\end{aligned}$$

implying

$$\epsilon > \psi(\tilde{y}, \tilde{x}, \boldsymbol{u}(\tilde{x})) - \psi(y, \tilde{x}, \boldsymbol{u}(\tilde{x})) \geq \boldsymbol{v}(\tilde{y}) - \boldsymbol{v}(y) \geq \psi(\tilde{y}, x, \boldsymbol{u}(x)) - \psi(y, x, \boldsymbol{u}(x)) > -\epsilon,$$

where the outer inequalities are from (24) and the fact that $\underline{u} \leq \boldsymbol{u}(x) \leq \bar{u}$ holds for all $x \in X$. Consequently, we have $\|\boldsymbol{v}(\tilde{y}) - \boldsymbol{v}(y)\| < \epsilon$, thus establishing (23). ■

4 Optimal Tariffs in the Principal-Agent Model

This section applies our characterizations of implementable profiles and pairs to establish the existence of optimal tariffs in a broad class of adverse-selection principal-agent models. We begin in Section 4.1 by introducing the principal-agent model and the principal's optimization problem. Section 4.2 reformulates the principal's optimization problem as the choice over a subset of *implementable* tariffs, which leads to a straightforward existence proof, presented in Section 4.3. Finally, we discuss the question of whether the optimal tariff will implement a utility profile satisfying the agent's participation constraint with equality for some type. Section 4.4 provides a sufficient condition for the participation constraint to be binding in this sense.

Carlier (2001), Carlier (2002), Jullien (2000), and Kahn (1993) have existence results for special cases of the principal-agent model with quasilinear utility. As Kadan, Reny, and Swinkels (2011) explain, the challenge in establishing a general existence result is to find a topology in which the set of objects to be implemented is compact and the principal's payoff is continuous. The taxation principle allows us to think of the principal as choosing a tariff, but allowing the principal to choose from the set of all tariffs leads to a stubborn tension between continuity and compactness. Balder (1996), Kadan, Reny, and Swinkels (2011), and Page (1991, 1992, 1997) respond to this difficulty by essentially expanding the set of objects from which the principal can choose. Balder (1996) and Page (1991, 1992, 1997) allow the principal to choose tariff correspondences rather than functions, while also imposing compactness assumptions on their (quite general) models that we avoid.[16] Kadan, Reny, and Swinkels (2011) expand the principal's choices to include random mechanisms.[17] In contrast, our response is to restrict rather than expand the principal's set of alternatives. We use our strengthening of the taxation principle (Corollary 3) to formulate the principal's problem as the nonlinear pricing problem of choosing an implementable tariff subject to the agent's participation constraint. We then show that the principal's choice can be restricted to a bounded subset of these tariffs—intuitively, the agents' participation constraint implies that the principal cannot set the tariff too high, while tariffs that are too low raise too little revenue. We then use the equicontinuity of bounded sets of implementable profiles (Corollary 5) to obtain compactness, opening the door to obtain the existence of optimal tariffs from a straightforward application of Weierstrass' extreme value theorem.

4.1 The Principal-Agent Model

Throughout the following we impose Assumption 1, interpreting $\phi(x, y, v)$ as the utility of an agent of type x, who takes decision y and provides transfer v to the principal. The corresponding utility of the principal is given by the function $\pi : X \times Y \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$, satisfying a regularity condition analogous to the one for the agent's utility function.

Assumption 2. *The function π is continuous, strictly increasing in its third argument, and satisfies $\pi(x, y, \overline{\mathbb{R}}) = \overline{\mathbb{R}}$ for all $(x, y) \in X \times Y$.*

Let $\underline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ be a profile of reservation utilities for the agents and let μ be a Borel probability measure on X, describing the principal's belief over the agent's type.

Using the strengthening of the taxation principle in Corollary 3 we may view the principal as choosing a tariff $\boldsymbol{v} \in \boldsymbol{I}(Y)$ and a pair $(\boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{I}(X) \times Y^X$ implemented by $\boldsymbol{v}$ subject to the agent's participation constraint, that is, $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$. Given any such

[16] Applying their assumptions to our model would restrict the possible values of the third argument of ϕ and ψ to be drawn from a compact set.

[17] We have followed much of the literature in assuming that the principal must offer deterministic contracts. As Strausz (2006) shows, this is restrictive. Like us, Kadan, Reny, and Swinkels (2011) do not require the third arguments of ϕ and ψ to be drawn from a compact set. In their setting this gives rise to formidable difficulties that do not appear in our analysis because we do not consider moral hazard.

choice, the resulting profit of the principal when faced with type x of the agent is $\pi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x))$. We thus formulate the principal's problem as

$$\max_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{y})\in \boldsymbol{I}(Y)\times\boldsymbol{I}(X)\times Y^X} \int_{x\in X} \pi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))d\mu(x) \tag{25}$$

subject to the constraints that (i) $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$, (ii) $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ holds, (iii) $z : X \to \mathbb{R}$, given by $z(x) = \pi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$ for all $x \in X$ is Borel measurable.[18] We say that $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{I}(Y) \times \boldsymbol{I}(X) \times Y^X$ is feasible in the principal's problem if it satisfies constraints (i)–(iii).

4.2 Reformulating the Principal's Problem

Unless the argmax-correspondence $\boldsymbol{Y_v}$ defined in (4) is single-valued, there are multiple decision assignments $\boldsymbol{y}$ implemented by a given tariff $\boldsymbol{v} \in \boldsymbol{I}(Y)$. To deal with this complication, we define the function $\hat{\pi} : X \times \boldsymbol{I}(Y) \to \mathbb{R}$ by

$$\hat{\pi}(x, \boldsymbol{v}) = \max_{y\in \boldsymbol{Y_v}(x)} \pi(x, y, \boldsymbol{v}(y)). \tag{26}$$

From Lemma 4.2 we have that $\boldsymbol{v} \in \boldsymbol{I}(Y)$ is continuous. As the generating function ϕ is continuous (Assumption 1), a standard application of Berge's maximum theorem (Ok, 2007, p.306) ensures that for all $\boldsymbol{v} \in \boldsymbol{I}(Y)$ the correspondence $\boldsymbol{Y_v} : X \rightrightarrows Y$ is non-empty-valued, compact-valued, and upper hemicontinuous. Hence, the correspondence $\hat{\boldsymbol{Y}} : X \times \boldsymbol{I}(Y) \rightrightarrows Y$ defined by

$$\hat{\boldsymbol{Y}}_{\boldsymbol{v}}(x) = \operatorname*{argmax}_{y\in \boldsymbol{Y_v}(x)} \pi(x, y, \boldsymbol{v}(y)) \tag{27}$$

is non-empty valued and the function $\hat{\pi}$ is well-defined and Borel measurable in x (Aliprantis and Border, 2006, Theorems 18.19 and 18.20). As every $\boldsymbol{y}$ implemented by $\boldsymbol{v}$ satisfies $\pi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)) \leq \hat{\pi}(x, \boldsymbol{v})$, any selection from $\hat{\boldsymbol{Y}}_{\boldsymbol{v}}$ satisfies this relationship with equality, and $\hat{\pi}$ is measurable in x it follows that we can rewrite the principal's problem as maximizing

$$\hat{\Pi}(\boldsymbol{v}) = \int_{x\in X} \hat{\pi}(x, \boldsymbol{v})d\mu(x)$$

over $(\boldsymbol{v}, \boldsymbol{u}) \in \boldsymbol{I}(Y) \times \boldsymbol{I}(X)$ subject to (i) the constraint that $\boldsymbol{v}$ implements $\boldsymbol{u}$ and (ii) the participation constraint $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$. As every tariff $\boldsymbol{v} \in \boldsymbol{I}(Y)$ is continuous (Lemma 4.2) and thus implements the rent function $\Phi\boldsymbol{v}$ (Lemma 3), we can replace these two

[18] Because X and Y are compact, $\boldsymbol{v} \in \boldsymbol{I}(Y)$ is bounded (Lemma 4.2), and π is continuous, measurability of z implies its integrability, thus ensuring that the objective function in (25) is well-defined on the feasible set of the maximization problem.

remaining constraints by $\Phi \boldsymbol{v} \geq \underline{\boldsymbol{u}}$. Using Proposition 1, we have $\Phi \boldsymbol{v} \geq \underline{\boldsymbol{u}} \iff \boldsymbol{v} \leq \Psi \underline{\boldsymbol{u}}$, yielding the following reformulation of the principal's problem

$$\max_{\boldsymbol{v} \in \boldsymbol{I}(Y)} \hat{\Pi}(\boldsymbol{v}) \text{ subject to } \boldsymbol{v} \leq \Psi \underline{\boldsymbol{u}}. \tag{28}$$

We refer to (28) as the nonlinear pricing problem as it reduces the principal's problem to the choice of an implementable tariff attaching a price to each decision $y \in Y$. Solving this nonlinear pricing problem yields a solution to the original formulation of the principal's problem. Specifically, if $\boldsymbol{v}$ is a solution to (28), then specifying $\boldsymbol{u} = \Phi \boldsymbol{v}$ and letting $\boldsymbol{y}$ be any selection from the correspondence $\hat{\boldsymbol{Y}}_{\boldsymbol{v}}$ defined in (27) provides a triple $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ solving (25) subject to the implementability, participation, and measurability constraints defining feasibility in the principal's problem.

4.3 Existence Result

We now establish the existence of a solution to the nonlinear pricing problem. We do so with the help of two lemmas which set the stage for an application of Weierstrass' extreme value theorem. First, we observe that the objective function $\hat{\Pi} : \boldsymbol{I}(Y) \to \mathbb{R}$ is upper semicontinuous. The proof is in Appendix A.6.

Lemma 6. *Let Assumptions 1 and 2 hold. Then* $\hat{\Pi}$ *is upper semicontinuous.*

Second, the set of tariffs that are feasible in the nonlinear pricing problem is bounded above by $\Psi \underline{\boldsymbol{u}} \in \boldsymbol{I}(Y)$. While there is no corresponding lower bound in the formulation of the nonlinear pricing problem, it is intuitive that a suitable lower bound can be imposed without impinging on the value of the principal's maximization problem. This is verified in the proof of the following lemma, given in Appendix A.7.

Lemma 7. *Let Assumptions 1 and 2 hold. Then there exists* $\underline{\boldsymbol{v}} \in \boldsymbol{I}(Y)$ *such that* $\hat{\Pi}(\boldsymbol{v}) \geq \hat{\Pi}(\Psi \underline{\boldsymbol{u}})$ *implies* $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$.

In conjunction with Corollaries 4 and 5, Lemma 7 implies that we can restrict the principal's choice set to a compact subset of the set of implementable tariffs. This yields:

Proposition 4. *Let Assumptions 1–2 hold. Then the nonlinear pricing problem* (28) *has a solution.*

Proof. To simplify notation, let $\overline{\boldsymbol{v}} = \Psi \underline{\boldsymbol{u}}$. By Lemma 7 there exists $\underline{\boldsymbol{v}} \in \boldsymbol{I}(Y)$ with the property that $\hat{\Pi}(\boldsymbol{v}) \geq \hat{\Pi}(\overline{\boldsymbol{v}})$ implies $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$. Fix such a $\underline{\boldsymbol{v}}$. Clearly, we have $\underline{\boldsymbol{v}} \leq \overline{\boldsymbol{v}}$. Thus, the order interval $[\underline{\boldsymbol{v}}, \overline{\boldsymbol{v}}] = \{\boldsymbol{v} \in \boldsymbol{B}(Y) | \underline{\boldsymbol{v}} \leq \boldsymbol{v} \leq \overline{\boldsymbol{v}}\}$ is a non-empty, closed, and bounded subset of $\boldsymbol{B}(Y)$. As $\boldsymbol{I}(Y)$ is also closed (Corollary 4), it follows that $\mathcal{V} = [\underline{\boldsymbol{v}}, \overline{\boldsymbol{v}}] \cap \boldsymbol{I}(Y)$ is a closed and bounded subset of $\boldsymbol{I}(Y)$. By Corollary 5 this set is also equicontinuous. Hence, the Arzela-Ascoli theorem (Ok, 2007, p. 264) implies that $\mathcal{V}$ is compact. As $\overline{\boldsymbol{v}}$ is an element of $\mathcal{V}$ this set is also non-empty. As $\hat{\Pi}$ is upper semicontinuous by Lemma 6, Weierstrass' extreme value theorem then implies that the problem

$$\max_{\boldsymbol{v}\in\boldsymbol{I}(Y)} \hat{\Pi}(v) \text{ subject to } \underline{\boldsymbol{v}} \leq \boldsymbol{v} \leq \overline{\boldsymbol{v}}$$

has a solution $\boldsymbol{v}^*$. As $\overline{\boldsymbol{v}}$ is in the feasible set of this problem, we have $\hat{\Pi}(\boldsymbol{v}^*) \geq \hat{\Pi}(\overline{\boldsymbol{v}})$ and thus (using the definition of $\underline{\boldsymbol{v}}$) $\hat{\Pi}(\boldsymbol{v}^*) \geq \hat{\Pi}(\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{I}(Y)$ satisfying $\boldsymbol{v} \leq \overline{\boldsymbol{v}} = \Psi\underline{\boldsymbol{u}}$. Hence, $\boldsymbol{v}^*$ solves the nonlinear pricing problem (28). ■

4.4 Is the Participation Constraint Binding?

Let $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ be a solution to the principal's problem and suppose the reservation utility profile $\underline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ is continuous. In the quasilinear case it is then straightforward that the participation constraint $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ is binding in the sense that there exists $x \in X$ satisfying $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$. For suppose the participation constraint is not binding. Then the principal can add a small constant payment to the tariff without interfering with the participation constraint. Because of quasilinearity, adding a constant to the tariff leaves the optimal choices of the agent types unaffected and thus results in a strictly higher expected payoff for the principal.

In the absence of quasilinearity, the situation is more difficult. The following example shows that the principal's problem may have a unique solution $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ satisfying $\boldsymbol{u}(x) > \underline{\boldsymbol{u}}(x)$ for all $x \in X$, so that the participation constraint is not binding.

Example 1. Let $X = \{1, 2\}$ and $Y = \{1, 2\}$. The two possible types are equally likely: $\mu(1) = \mu(2) = 1/2$. The reservation utility profile is given by $\underline{\boldsymbol{u}}(1) = \underline{\boldsymbol{u}}(2) = 0$. The agent's utility function is

$$\begin{aligned}
\phi(1,1,v) &= 3 - 2v \\
\phi(1,2,v) &= 2 - v \\
\phi(2,1,v) &= \frac{3}{2} - \frac{1}{2}v \\
\phi(2,2,v) &= 2 - v
\end{aligned}$$

and the principal's utility function is

$$\begin{aligned}
\pi(1,1,v) &= v + 5 \\
\pi(1,2,v) &= v \\
\pi(2,1,v) &= v \\
\pi(2,2,v) &= v + 5.
\end{aligned}$$

Let $(\boldsymbol{v}^*, \boldsymbol{u}^*, \boldsymbol{y}^*)$ be given by $\boldsymbol{v}^*(1) = \boldsymbol{v}^*(2) = 1$, $\boldsymbol{u}^*(1) = \boldsymbol{u}^*(2) = 1$, and $\boldsymbol{y}^*(1) = 1$, $\boldsymbol{y}^*(2) = 2$. It is clear that $(\boldsymbol{v}^*, \boldsymbol{u}^*, \boldsymbol{y}^*)$ satisfies the feasibility constraints of the principal's problem, that the participation constraint is not binding, and that the principal's payoff is $\Pi(\boldsymbol{v}^*) = 6$. In the following we will argue that $(\boldsymbol{v}^*, \boldsymbol{u}^*, \boldsymbol{y}^*)$ is the unique solution to the principal's problem. To do so, it suffices to show that there is no other feasible $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ satisfying $\Pi(\boldsymbol{v}) \geq 6$.

Consider, first, any feasible $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ with $\boldsymbol{y} \neq \boldsymbol{y}^*$. The participation constraint implies $\boldsymbol{v}(1) \leq 2$ and $\boldsymbol{v}(2) \leq 3$, so that (because $\boldsymbol{y}(1) \neq 1$ or $\boldsymbol{y}(2) \neq 2$) the principal's payoff is bounded above by 5.

Consider, second, any feasible $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ with $\boldsymbol{y} = \boldsymbol{y}^*$. If we are to implement $\boldsymbol{y}^*$, agent 1 must prefer decision 1 to decision 2, or $3 - 2\boldsymbol{v}(1) \geq 2 - \boldsymbol{v}(2)$, which is

$$\boldsymbol{v}(2) \geq 2\boldsymbol{v}(1) - 1. \tag{29}$$

Similarly, agent 2 must prefer decision 2 to decision 1, or $2 - \boldsymbol{v}(2) \geq \frac{3}{2} - \frac{1}{2}\boldsymbol{v}(1)$, which is

$$2\boldsymbol{v}(2) \leq 1 + \boldsymbol{v}(1). \tag{30}$$

A necessary condition for (29)–(30) is that $\boldsymbol{v}(1) \leq 1$ and $\boldsymbol{v}(2) \leq 1$. As the principal's utility is strictly increasing in the transfer v, it follows that $\boldsymbol{v}^*(1) = \boldsymbol{v}^*(2) = 1$ is the unique optimal tariff. ■

The key feature of Example 1 is that it is simply impossible to implement the optimal decision assignment $\boldsymbol{y}$ while satisfying the participation constraint of some type of agent with equality. A straightforward variation on this example shows that there may exist assignments that are implementable, but only if a participation constraint is violated. In the following we provide a condition sufficient to exclude these possibilities.

Let us say that an assignment $\boldsymbol{y} \in Y^X$ is *strongly implementable* if for every $(x_0, r_0) \in X \times \mathbb{R}$ there exists $\boldsymbol{u} \in \boldsymbol{B}(X)$ satisfying $\boldsymbol{u}(x_0) = r_0$ such that $(\boldsymbol{u}, \boldsymbol{y})$ is implementable. Strong implementability of $\boldsymbol{y}$ is thus the requirement that $\boldsymbol{y}$ can be implemented with any initial condition (x_0, r_0) specifying a target utility level $u(x_0) = r_0$ for type $x_0 \in X$.[19] It is immediate that in the quasilinear case every implementable decision profile $\boldsymbol{y}$ is strongly implementable: whenever $\boldsymbol{v}$ implements $\boldsymbol{y}$, the tariff $\tilde{\boldsymbol{v}}$ defined by $\tilde{\boldsymbol{v}}(y) = \boldsymbol{v}(y) + [r_0 - \boldsymbol{v}(\boldsymbol{y}(x_0))]$ for all $y \in Y$ will also implement $\boldsymbol{y}$ and satisfies the initial condition (x_0, r_0). As we demonstrate in Section 5.3, it is also the case that all implementable decision assignments are strongly implementable whenever the generating function ϕ satisfies a strict single crossing property. Hence, the following result is applicable beyond the quasilinear case (and, as we will see in Section 6, of relevance not only for the question we study here, but also for the existence of stable matchings).

[19] Of course, the notion of a strongly implementable $\boldsymbol{x} \in X^Y$ can be defined in an analogous way and results dual to the ones we give below will hold for such assignments.

Proposition 5. *Let Assumption 1 hold. For every strongly implementable $\boldsymbol{y} \in Y^X$ and every $\underline{\boldsymbol{u}} \in \boldsymbol{CB}(X)$ there exists $\boldsymbol{u} \in \boldsymbol{B}(X)$ such that $(\boldsymbol{u}, \boldsymbol{y})$ is implementable, $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ holds, and there exists $x \in X$ satisfying $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$.*

Establishing this result is non-trivial: While strong implementability of $\boldsymbol{y}$ guarantees that for every $x \in X$ there exists $\boldsymbol{u}$ such that $(\boldsymbol{u}, \boldsymbol{y})$ is implementable and satisfies $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$, it is not obvious that any such $\boldsymbol{u}$ satisfies the participation constraint $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$. We tackle this issue by first exploiting the duality results from Section 3 to show that the set of tariffs implementing a given assignment has a lattice structure and that this structure is inherited by the set of implemented rent functions. Appendix A.8 proves:[20]

Lemma 8. *Let Assumption 1 hold. Suppose $\boldsymbol{v}_1$ implements $(\boldsymbol{u}_1, \boldsymbol{y})$ and that $\boldsymbol{v}_2$ implements $(\boldsymbol{u}_2, \boldsymbol{y})$. Then $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$ implements $(\boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{y})$ and $\boldsymbol{v}_1 \vee \boldsymbol{v}_2$ implements $(\boldsymbol{u}_1 \wedge \boldsymbol{u}_2, \boldsymbol{y})$.*

Building on Lemma 8, the proof of Proposition 5 (provided in Appendix A.9) constructs a sequence of tariffs implementing $\boldsymbol{y}$ such that the corresponding sequence of implementable rent functions converges to an implementable rent function with the property that the participation constraint is satisfied and holds with equality for some type.

Returning to the question which motivated our consideration of strongly implementable profiles, we can now state conditions that suffice for the participation constraint to be binding in a solution to the principal's problem.

Corollary 6. *Let Assumptions 1 and 2 hold and suppose further that the reservation utility profile satisfies $\underline{\boldsymbol{u}} \in \boldsymbol{CB}(X)$ and the Borel probability measure μ is strictly positive.*[21] *If $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ solves the principal's problem and $\boldsymbol{y}$ is strongly implementable, then there exists $x \in X$ such that $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$ holds.*

Proof. Suppose $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ is feasible in the principal's problem, $\boldsymbol{y}$ is strongly implementable, and $\boldsymbol{u} > \underline{\boldsymbol{u}}$ holds. By Proposition 5 there exists $\boldsymbol{v}' \in \boldsymbol{B}(Y)$ implementing $(\boldsymbol{y}, \boldsymbol{u}')$ such that $\boldsymbol{u}' \geq \underline{\boldsymbol{u}}$ holds and there exists $\hat{x} \in X$ satisfying $\boldsymbol{u}'(\hat{x}) = \underline{\boldsymbol{u}}(\hat{x})$. Let $\tilde{\boldsymbol{v}} = \boldsymbol{v} \vee \boldsymbol{v}'$ and $\tilde{\boldsymbol{u}} = \boldsymbol{u} \wedge \boldsymbol{u}'$. By Lemma 8, $\tilde{\boldsymbol{v}}$ implements $(\boldsymbol{y}, \tilde{\boldsymbol{u}})$. Because $\boldsymbol{u}$ and $\boldsymbol{u}'$

[20] The last part of this result, namely that the lattice structure is inherited by the set of implemented rent functions, can also be obtained as an implication of a more general result (Davey and Priestley, 2002, Proposition 7.31): because the implementation maps are a Galois connection (Proposition 1) the implementation maps reverse all existing (and not only pairwise) meets and joins. We note that (in conjunction with Proposition 2) this general result implies that the sets of implementable profiles $\boldsymbol{I}(X)$ and $\boldsymbol{I}(Y)$ are conditionally complete sublattices of $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$.

[21] As suggested by the discussion at the beginning of this subsection, in the quasilinear case the assumption that μ is strictly positive is not required because in this case the set S appearing in the following proof is equal to X and the assumption is only required to ensure that S has strictly positive measure.

both satisfy the participation constraint, so does $\tilde{\boldsymbol{u}}$. Further, we have $\boldsymbol{u} \geq \tilde{\boldsymbol{u}}$ with strict inequality holding on a non-empty open set $S \subset X$ (because strict inequality holds at $\hat{x} \in X$ and both $\boldsymbol{u}$ and $\tilde{\boldsymbol{u}}$ are continuous by Lemma 4.2.). As ϕ is strictly increasing in its third argument and both $\boldsymbol{u}$ and $\tilde{\boldsymbol{u}}$ implement $\boldsymbol{y}$ we can not only infer (from the order reversal property in Corollary 1.2) the inequality $\tilde{\boldsymbol{v}} \geq \boldsymbol{v}$, but also that $\tilde{\boldsymbol{v}}(\boldsymbol{y}(x)) > \boldsymbol{v}(\boldsymbol{y}(x))$ holds for all $x \in S$. Using Assumption 2 this implies $\pi(x, \boldsymbol{y}(x), \tilde{\boldsymbol{v}}(\boldsymbol{y}(x))) \geq \pi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$ for all $x \in X$ with strict inequality holding for $x \in S$. Now let $\tilde{\boldsymbol{y}}$ be any selection from the argmax-correspondence $\hat{\boldsymbol{Y}}_{\boldsymbol{v}}$ of the problem defining $\hat{\pi}(x, \tilde{\boldsymbol{v}})$ (cf. (27)). We then have that $(\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{y}}, \tilde{\boldsymbol{u}})$ is feasible in the principal's problem. Further, by definition of $\tilde{\boldsymbol{y}}$ we have $\pi(x, \tilde{\boldsymbol{y}}(x), \tilde{\boldsymbol{v}}(\tilde{\boldsymbol{y}}(x))) \geq \pi(x, \boldsymbol{y}(x), \tilde{\boldsymbol{v}}(\boldsymbol{y}(x)))$ and thus $\pi(x, \tilde{\boldsymbol{y}}(x), \tilde{\boldsymbol{v}}(\tilde{\boldsymbol{y}}(x))) \geq \pi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$ for all $x \in X$ with strict inequality holding for $x \in S$. Using strict positivity of μ this implies $\hat{\Pi}(\tilde{\boldsymbol{v}}) > \hat{\Pi}(\boldsymbol{v})$. Hence, $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ satisfying $\boldsymbol{u} > \underline{\boldsymbol{u}}$ cannot be a solution to the principal's problem. ■

5 Implementable Assignments

This section examines the set of implementable assignments. We present our analysis in terms of implementing an assignment $\boldsymbol{y}$ and use the terminology from the principal-agent model, but of course a dual argument establishes the corresponding results for assignments $\boldsymbol{x}$.

As we have explained in the Section 4.4, in the quasilinear case one need not worry about initial conditions (specifying a utility level for some type of the agent) when considering the implementability of an assignment: one can first ask whether *any* tariff implements a given decision assignment, and then can add or subtract a constant from this tariff to adjust initial conditions at will while still implementing the desired assignment. Without quasilinearity, adding a constant to a tariff in general changes the assignment implemented by the tariff. Initial conditions and implementability must then be considered together. In particular, as illustrated by Example 1 in Section 4.4, a decision assignment may be implementable for some initial conditions, but may fail to be strongly implementable.

The main result of this section is that in the presence of a single-crossing condition (Assumption 3) all increasing assignments are strongly implementable (Proposition 7). When the single-crossing condition holds strictly, this yields a complete characterization of implementable allocations. Here, as in the quasilinear case, an assignment is either implementable for all or no initial conditions and it is implementable if and only if it is increasing. These results obtain whenever the generating function satisfies Assumption 1. We thus generalize a corresponding characterization result of Guesnerie and Laffont (1994, Corollary 2.1) by dispensing with the smoothness conditions on the generating function and the restriction to piecewise continuously differentiable assignments maintained by these authors.

In the quasilinear case it is obvious that the arguments given in the proof of

Proposition 1 in Rochet (1987) carry over without any smoothness assumptions to establish that under the strict single-crossing condition an assignment is strongly implementable if and only if it is increasing. In contrast, the arguments given for the general case in Guesnerie and Laffont (1994) rely on the existence of the solution to a differential equation and have no obvious counterpart when the agent's utility function is only required to be continuous. We thus begin the analysis in Section 5.1 by building on the results from Section 3 to develop a radically different approach to the characterization of implementable assignments. Section 5.2 provides a sufficient condition ensuring the implementability of a pair $(\boldsymbol{y}, \boldsymbol{u})$ with given initial condition (Proposition 6). This result requires nothing but Assumption 1. Section 5.3 then introduces the single-crossing condition and shows that under this condition every increasing decision assignment satisfies the sufficient condition from Proposition 6 for every initial condition. This yields the strong implementability results under the (strict) single-crossing condition mentioned above.

5.1 A Fixed Point Characterization

We consider a given $\boldsymbol{y} \in Y^X$ and a given $(x_0, r_0) \in X \times \mathbb{R}$ and address the question whether the assignment $\boldsymbol{y}$ is implementable with initial condition (x_0, r_0), i.e., whether there exists $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$ such that the profile $\boldsymbol{v}$ implements the pair $(\boldsymbol{y}, \boldsymbol{u})$ and the profile $\boldsymbol{u}$ satisfies $\boldsymbol{u}(x_0) = r_0$. From Lemma 4.2 and Corollary 2 we may limit our attention to $\boldsymbol{v} \in \boldsymbol{CB}(Y)$.

Suppose $\boldsymbol{v}$ implements $(\boldsymbol{y}, \boldsymbol{u})$ and let $y_0 = \boldsymbol{y}(x_0)$. We then have $\boldsymbol{u}(x_0) = \Phi\boldsymbol{v}(x_0) = \phi(x_0, y_0, \boldsymbol{v}(y_0))$. Hence, $\boldsymbol{u}$ will satisfy the initial condition (x_0, r_0) if and only if

$$\phi(x_0, y_0, \boldsymbol{v}(y_0)) = r_0,$$

or, equivalently,

$$\boldsymbol{v}(y_0) = \psi(y_0, x_0, r_0),$$

where the equivalence is from (1). Hence, if we let $v_0 = \psi(y_0, x_0, r_0)$ the requirement that $\boldsymbol{v}$ implements $\boldsymbol{y}$ with initial condition (x_0, r_0) is equivalent to the requirement that $\boldsymbol{v}$ implements $\boldsymbol{y}$ and satisfies $\boldsymbol{v}(y_0) = v_0$. Letting $\boldsymbol{CB}_0(Y) \subset \boldsymbol{CB}(Y)$ be the set of continuous and bounded profiles $\boldsymbol{v}$ satisfying $\boldsymbol{v}(y_0) = v_0$, we can thus focus on the question whether there is $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ implementing $\boldsymbol{y}$ throughout the following and we will proceed accordingly.

Define $\tilde{\Phi} : \boldsymbol{CB}_0(Y) \to \boldsymbol{B}(X)$ by

$$\tilde{\Phi}\boldsymbol{v}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x))) \quad \forall x \in X, \tag{31}$$

so that $\tilde{\Phi}\boldsymbol{v}$ is the utility profile which results if all types x of the agent are forced to take the decision $\boldsymbol{y}(x)$ and pay the resulting transfer $\boldsymbol{v}(\boldsymbol{y}(x))$. Now suppose that $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ implements $\boldsymbol{y}$. Then choosing $\boldsymbol{y}(x)$ when faced with the tariff $\boldsymbol{v}$ is optimal for all types of the agents, yielding the equality $\tilde{\Phi}\boldsymbol{v} = \Phi\boldsymbol{v}$. Applying the implementation map Ψ to this equality and noting that, from Proposition 3, we may take $\boldsymbol{v}$ to be implementable and thus satisfy $\boldsymbol{v} = \Psi\Phi\boldsymbol{v}$ (Proposition 2), we obtain that the existence of $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ satisfying the fixed point condition $\boldsymbol{v} = \Psi\tilde{\Phi}\boldsymbol{v}$ is necessary for $\boldsymbol{y}$ to be implementable with initial condition (x_0, r_0). The proof of the following lemma shows that this condition is also sufficient, yielding:

Lemma 9. *Let Assumption 1 hold. Assignment $\boldsymbol{y} \in Y^X$ is implementable with initial condition (x_0, r_0) if and only if there exists $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ satisfying*

$$\boldsymbol{v} = \Psi\tilde{\Phi}\boldsymbol{v}. \tag{32}$$

Proof. As the argument before the statement of the lemma has established necessity, it remains to show that $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ satisfying (32) implements $\boldsymbol{y}$. Thus, suppose $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ satisfies (32). Because $\boldsymbol{v}$ is continuous, it suffices to show that this implies $\tilde{\Phi}\boldsymbol{v} = \Phi\boldsymbol{v}$ as then $\phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x))) = \max_{y \in Y} \phi(x, y, \boldsymbol{v}(y))$ holds for all $x \in X$. The inequality $\tilde{\Phi}\boldsymbol{v} \leq \Phi\boldsymbol{v}$ is immediate from the definition of the map $\tilde{\Phi}$. To obtain the reverse inequality, apply the cancellation property (12) from Corollary 1.3 to the profile $\tilde{u} = \tilde{\Phi}\boldsymbol{v}$ to obtain $\tilde{\Phi}\boldsymbol{v} \geq \Phi\Psi\tilde{\Phi}\boldsymbol{v}$. Using (32), the right side of this inequality is equal to $\Phi\boldsymbol{v}$, so that we obtain $\tilde{\Phi}\boldsymbol{v} \geq \Phi\boldsymbol{v}$. ■

Condition (32) in the statement of Lemma 9 is similar in form to the fixed point characterization of implementable profiles $\boldsymbol{v}$ appearing in Proposition 2.2. Indeed, because Ψ maps onto the set of implementable profiles $\boldsymbol{I}(Y)$, condition (32) characterizes the set of implementable profiles $\boldsymbol{v}$ which implement the assignment $\boldsymbol{y}$.

5.2 A Sufficient Condition

Lemma 9 directs our attention to identifying properties of the assignment $\boldsymbol{y}$ and the initial condition (x_0, r_0) which ensure that the mapping $\Psi\tilde{\Phi}$ has a fixed point. As a first step in this direction, we demonstrate in Lemma 10 that a modified version of the mapping $\Psi\tilde{\Phi}$ always has a fixed point. In a second step, we then provide a condition which ensures that the fixed points of the modified mapping satisfy the fixed point condition (32), thus yielding the sufficient condition for the implementability of $\boldsymbol{y}$ with initial condition (x_0, r_0) recorded in Proposition 6 below.

We introduce two modifications of the map $\Psi\tilde{\Phi}$. The first modification replaces the map $\tilde{\Phi}$ by a map $\hat{\Phi}$ in such a way that the image of $\boldsymbol{CB}_0(Y)$ under the map $\Psi\hat{\Phi}$ is bounded, while not interfering with the fixed point characterization in Lemma 9. The second modification subtracts a constant from the output of the map $\Psi\hat{\Phi}$ to obtain a self-map of the set $\boldsymbol{CB}_0(Y)$.

In the profile $\tilde{\Phi}\boldsymbol{v}$ agent x_0 receives utility r_0. The ability to assign agent x_0 to any decision y thus imposes a lower bound on $\Psi\tilde{\Phi}(\boldsymbol{v})(y)$. Our first modification is designed to ensure that there exist a corresponding upper bound. Fix $\varepsilon > 0$ and define $\underline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ by

$$\underline{\boldsymbol{u}}(x) = \phi(x, y_0, v_0) - \varepsilon \tag{33}$$

and then define $\hat{\Phi} : \boldsymbol{CB}_0(Y) \to \boldsymbol{B}(X)$ by setting

$$\hat{\Phi}\boldsymbol{v} = \tilde{\Phi}\boldsymbol{v} \vee \underline{\boldsymbol{u}}. \tag{34}$$

We can then interpret $\hat{\Phi}\boldsymbol{v}(x)$ as the maximum of the utilities agent x can secure by either accepting to pay $\boldsymbol{v}(\boldsymbol{y}(x))$ for decision $\boldsymbol{y}(x)$ or by accepting an outside option resulting in utility $\underline{\boldsymbol{u}}(x)$ (with the ε in (33) ensuring that indifferences are broken the right way). Provided that the inequality $\phi(x, \boldsymbol{y}, \boldsymbol{v}(\boldsymbol{y}(x))) \geq \phi(x, y_0, v_0)$ holds, it follows from (33) that agent x will prefer to pay $\boldsymbol{v}(\boldsymbol{y}(x))$ for decision $\boldsymbol{y}(x)$, so that we have $\hat{\Phi}\boldsymbol{v}(x) = \tilde{\Phi}\boldsymbol{v}(x)$. In particular, this equality always holds for type x_0.

For the second modification, define $\rho : \boldsymbol{CB}(Y) \to \boldsymbol{CB}_0(Y)$ by

$$\rho\boldsymbol{v}(y) = \boldsymbol{v}(y) - [\boldsymbol{v}(y_0) - v_0] \tag{35}$$

and then let $T : \boldsymbol{CB}_0(Y) \to \boldsymbol{CB}_0(Y)$ be given by

$$T\boldsymbol{v} = \rho\Psi\hat{\Phi}\boldsymbol{v}. \tag{36}$$

The mapping T thus subtracts for all $y \in Y$ the constant $\hat{\boldsymbol{v}}(y_0) - v_0$ from the profile $\hat{\boldsymbol{v}} = \Psi\hat{\Phi}\boldsymbol{v}$. (By Lemma 4, $\hat{\boldsymbol{v}} = \Psi\hat{\Phi}\boldsymbol{v}$ is an element of $\boldsymbol{CB}(Y)$, ensuring that T does indeed map into this set). This implies that the resulting tariff $T\hat{\boldsymbol{v}}$ satisfies the condition $T\hat{\boldsymbol{v}}(y_0) = v_0$ and hence is an element of $\boldsymbol{CB}_0(Y)$.

Making use of our results in Section 3 we can apply Schauder's fixed point theorem to show that the mapping T possesses a fixed point. Appendix A.10 proves:

Lemma 10. *Let Assumption 1 hold. There exists $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ satisfying*

$$\boldsymbol{v} = T\boldsymbol{v}. \tag{37}$$

Given the characterization result in Lemma 9 and the existence result in Lemma 10, it is immediate that any condition which yields the implication $\boldsymbol{v} = T\boldsymbol{v} \Longrightarrow \boldsymbol{v} = \Psi\tilde{\Phi}\boldsymbol{v}$ is a sufficient condition for the implementability of $\boldsymbol{y}$ with initial condition (x_0, r_0). The following proposition provides one such condition. The usefulness of this condition will be demonstrated in the following subsection. Appendix A.11 proves:

Proposition 6. *Let Assumption 1 hold. Suppose for every $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ there exists $\tilde{y} \in Y$ such that*

$$\boldsymbol{v}(\tilde{y}) = \Psi\tilde{\Phi}\boldsymbol{v}(\tilde{y}). \tag{38}$$

Then $\boldsymbol{y}$ is implementable with initial condition (x_0, r_0).

Condition (38) is again a fixed point condition and at first glance may appear similar to condition (32) in Lemma 9. The spirit is quite different, however: Rather than requiring that there exists some tariff $\boldsymbol{v}$ which stays fixed under the mapping $\Psi\tilde{\Phi}$, the requirement now is that for any tariff $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ there exists some decision $\tilde{y}$ such that the mapping $\Psi\tilde{\Phi}$ leaves the transfer $\boldsymbol{v}(\tilde{y})$ unchanged. The proof of Proposition 6 shows that under (38), any fixed point $\boldsymbol{v}$ of the mapping $T = \rho\Psi\hat{\Phi}$ satisfies (32), i.e., is also a fixed point of $\Psi\tilde{\Phi}$.

5.3 The Single-Crossing Condition

In the most commonly studied versions of the principal-agent model and the matching model the sets X and Y are unidimensional and the generating function satisfies a single crossing condition:

Assumption 3 (Single Crossing Condition)**.** *$X \subset \mathbb{R}$, $Y \subset \mathbb{R}$, and*

$$\phi(x_1, y_2, v_2) \geq \phi(x_1, y_1, v_1) \Rightarrow \phi(x_2, y_2, v_2) \geq \phi(x_2, y_1, v_1) \tag{39}$$

holds for all $x_1 < x_2 \in X$, $y_1 < y_2 \in Y$, and $v_1, v_2 \in \mathbb{R}$.

Note that Assumption 3 does not require X and Y to be intervals. Remark 4 below explains how the assumption $X \subset \mathbb{R}$ and $Y \subset \mathbb{R}$ can be relaxed.

If the second inequality in (39) holds as a strict inequality, then we say that ϕ satisfies the *strict single crossing condition.*[22] In the quasilinear case the (strict) single crossing condition is equivalent to requiring unidimensionality of X and Y and (strict) supermodularity of the function f.

Provided that the argmax-correspondence $\boldsymbol{Y_v}$ is nonempty-valued, the (strict) single crossing condition ensures the existence of increasing selections from this correspondence (resp. that all such selections must be increasing). Importantly for our subsequent arguments, the single crossing condition on ϕ implies an analogous single crossing condition on the inverse generating function ψ, ensuring that the argmax-correspondence $\boldsymbol{X_u}$ possesses analogous monotonicity properties. Using standard tools from the literature on monotone comparative statics, Appendix A.12 proves:

Lemma 11. *Let Assumptions 1 and 3 hold and let $\boldsymbol{v} \in \boldsymbol{B}(Y)$ be lower semicontinuous. Then there exists an increasing assignment $\boldsymbol{y}$ implemented by $\boldsymbol{v}$. Further, if ϕ satisfies the strict single crossing condition, then every assignment implemented by $\boldsymbol{v}$ is increasing. Dually, let $\boldsymbol{u} \in \boldsymbol{B}(X)$ be lower semicontinuous. Then there exists an increasing assignment $\boldsymbol{x}$ implemented by $\boldsymbol{u}$. Further, if ϕ satisfies the strict single crossing condition, then every assignment implemented by $\boldsymbol{u}$ is increasing.*

[22] In the context of a matching model, Legros and Newman (2007a, Section 4.2) say that ϕ satisfies (strict) generalized increasing differences if it satisfies a condition equivalent to the (strict) single crossing condition.

Equipped with the sufficient condition for the implementability of an assignment from Proposition 6 and with Lemma 11 the proof of the following is straightforward for the case of continuous decision assignments. The proof for the general case is more cumbersome and hence relegated to Appendix A.13.

Proposition 7. *Let Assumptions 1 and 3 hold. Every increasing assignment $\boldsymbol{y} \in Y^X$ is strongly implementable.*

Proof. Fix an increasing $\boldsymbol{y} \in Y^X$ and an initial condition $(x_0, r_0) \in \mathbb{R}^2$. Let $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$. It suffices to show that there exists $\tilde{y} \in Y$ such that (38) holds.

Suppose $\boldsymbol{y}$ is continuous. Then, because $\boldsymbol{v}$ and ϕ are continuous, so is $\tilde{\boldsymbol{u}} = \tilde{\Phi}\boldsymbol{v}$. From Lemma 11 there then exists an increasing $\boldsymbol{x} \in X^Y$ implemented by $\tilde{\boldsymbol{u}}$, that is

$$\Psi\tilde{\Phi}\boldsymbol{v}(y) = \psi(y, \boldsymbol{x}(y), \tilde{\boldsymbol{u}}(\boldsymbol{x}(y))) \tag{40}$$

holds for all $y \in Y$. By definition of the mapping $\tilde{\Phi}$ we also have

$$\tilde{\boldsymbol{u}}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x))) \tag{41}$$

for all $x \in X$. Now consider the mapping from $X \times Y$ to $X \times Y$ defined by $(x, y) \to (\boldsymbol{x}(y), \boldsymbol{y}(x))$. As X and Y are complete lattices and the functions $\boldsymbol{x}$ and $\boldsymbol{y}$ are both increasing, Tarski's fixed point theorem implies that this mapping has a fixed point $(\tilde{x}, \tilde{y}) \in X \times Y$. Using (40)–(41), for such a fixed point we have:

$$\Psi\tilde{\Phi}\boldsymbol{v}(\tilde{y}) = \psi(\tilde{y}, \tilde{x}, \tilde{\boldsymbol{u}}(\tilde{x})) \tag{42}$$

$$\tilde{\boldsymbol{u}}(\tilde{x}) = \phi(\tilde{x}, \tilde{y}, \boldsymbol{v}(\tilde{y})). \tag{43}$$

Substituting the expression for $\tilde{\boldsymbol{u}}(\tilde{x})$ from (43) into (42) and applying the inverse relationship (2) we obtain $\Psi\tilde{\Phi}\boldsymbol{v}(\tilde{y}) = \boldsymbol{v}(\tilde{y})$, which is the desired result.

If $\boldsymbol{y}$ is not continuous, we have no guarantee that the profile $\tilde{\boldsymbol{u}}$ defined by $\tilde{\boldsymbol{u}}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$ is lower semicontinuous and, hence, no guarantee that $\tilde{\boldsymbol{u}}$ implements an increasing assignment $\boldsymbol{x} \in X^Y$, disrupting a key step in the above proof. Replacing $\tilde{\boldsymbol{u}}$ in the above argument by its lower semicontinuous hull (cf. the proof of Lemma 4) fixes this difficulty. The details are spelled out in Appendix A.13. ∎

We note two corollaries to Proposition 7. First, under the strict single-crossing condition we have a complete characterization of (strongly) implementable assignments.

Corollary 7. *Let Assumptions 1 and 3 hold and suppose ϕ satisfies the strict single-crossing condition. Then an assignment $\boldsymbol{y} \in Y^X$ is implementable if and only if if is increasing and thus implementable if and only it it is strongly implementable.*

Proof. From Proposition 7, every increasing assignment $\boldsymbol{y} \in Y^X$ is strongly implementable (and thus implementable). From Proposition 3 and Lemma 4.2 every implementable assignment is implemented by a continuous profile. Hence, Lemma 11 implies that every implementable assignment is increasing (and thus strongly implementable). ■

Second, under the conditions in the statement of Corollary 7, every implementable assignment is strongly implementable so that every solution to the principal's problem considered in Section 4 features a strongly implementable decision assignment $\boldsymbol{y}$. The following is then immediate from Corollary 6.[23]

Corollary 8. *Let Assumptions 1–3 hold and suppose ϕ satisfies the strict single crossing condition. If the Borel measure μ is strictly positive and the reservation utility profile $\underline{\boldsymbol{u}}$ is continuous, then the participation constraint is binding in every solution to the principal's problem.*

Remark 4 (Multidimensional Decisions)**.** For the case of continuous decision assignments, Proposition 7 continues to hold when the assumption $Y \subset \mathbb{R}$ is replaced by the assumption that Y is a complete lattice and the single-crossing condition is imposed on the inverse generating function ψ. These assumptions suffice to ensure (by arguments analogous to the ones proving Lemma 11) that for lower semicontinuous $\boldsymbol{u} \in \boldsymbol{B}(X)$ an increasing selection $\boldsymbol{x}$ from the argmax-correspondence $\boldsymbol{X_u}$ exists—which, for continuous $\boldsymbol{y}$, is the only implication of Assumption 3 used in the proof of Proposition 7. This provides a generalization of Theorem 2 in Guesnerie and Laffont (1994). For multidimensional Y strict single crossing is not sufficient to imply that implementable assignments are increasing, so there is no counterpart to Corollary 7 for this case. ■

Remark 5 (The Full Range Assumption)**.** To obtain their counterpart to Corollary 7, Guesnerie and Laffont (1994, Corollary 2.1) assume that the marginal rate of substitution between decisions and transfers eventually does not grow too rapidly (i.e., is bounded by a linear function of v) as transfers increase (c.f. Fudenberg and Tirole, 1991, Theorem 7.3).[24] It is intuitive that some such condition is needed

[23]Using differentiability and monotonicity assumptions, Hellwig (2010, Theorem 2.1) shows that (weak) single crossing suffices to ensure that if $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ solves the principal's problem, then there exists $(\boldsymbol{v}', \boldsymbol{y}')$ with $\boldsymbol{y}'$ increasing such that $(\boldsymbol{v}', \boldsymbol{u}, \boldsymbol{y}')$ also solves the principal's problem. Applying Proposition 7 and Corollary 6 to $(\boldsymbol{v}', \boldsymbol{u}, \boldsymbol{y}')$ shows that the participation constraint binds in this solution to the principal's problem and thus (because $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{y})$ features the same utility profile) in every solution to the principal's problem. Hence, under Hellwig's conditions the strict single crossing requirement in Corollary 8 can be replaced by the single crossing condition.

[24] In our notation the condition is that ϕ is twice continuously differentiable and that there exist constants K_0 and $K_1 > 0$ such that

$$\left| \frac{\partial \phi(x, y, v)/\partial y}{\partial \phi(x, y, v)/\partial v} \right| \leq K_0 + K_1 |v|$$

for all $(x, y) \in X \times Y$.

to ensure that all increasing assignments are strongly implementable: if marginal rates of substitution increase too rapidly, there are bounds on the effective "utility transfers" that the principal can make to high agents, and the principal may not be able to make transfers large enough to induce separation. Technically, in Guesnerie and Laffont (1994) the implementation problem is reduced to solving a differential equation, and the restriction on the marginal rates of substitution is used to ensure the existence of a solution to this equation. In our analysis the assumption that the generating function ϕ has full range, which we have imposed as part of Assumption 1, plays an analogous role.[25] Roth and Sotomayor (1990, p. 223) offer a discussion of a similar full-range assumption in the context of a matching model. ■

6 Stable Assortative Matching

This section establishes an existence result for matching models in which buyer and seller types are continuously distributed on intervals in the real numbers and the generating functions describing the utility possibilities available to a matched pair of agents satisfy the single crossing condition (39). We provide conditions under which stable, positive-assortative matchings exist, thus filling a gap in the analysis of Nöldeke and Samuelson (2014) by providing an existence result for the leading case of the matching model studied in Section 4.3 of that paper.[26]

A stable matching differs from an implementable assignment (cf. Section 5) in two respects. First, a stable matching must satisfy a feasibility requirement. One may be able to assign all agents to the same decision, but one cannot assign all buyers to the same seller. Second, the stable matching must satisfy participation constraints on both sides of the market. Addressing feasibility is straightforward.

[25]To compare these conditions in a transparent setting, let us assume that ϕ is additively separable, i.e., $\phi(x, y, v) = f(x, y) - h(v)$, where f is continuously differentiable and h is continuously differentiable and strictly increasing. Then, since X and Y are compact, the condition from Guesnerie and Laffont (1994) (stated in footnote 24 above) reduces to the existence of K_0 and K_1 with $(h'(v))^{-1} \leq K_0 + K_1|v|$. If this condition is satisfied, then for $v \geq 0$ we have $h(v) = \int_0^v h'(\tau)d\tau + h(0) \geq \int_0^t (K_0 + K_1\tau)^{-1}d\tau + h(0) \geq (K_1)^{-1}\ln(K_0 + K_1 v) + h(0)$, which (with a similar argument for $v \leq 0$) ensures that h has full range. The example of the function

$$h(v) = \begin{cases} \sqrt{\ln(1+v)} & v \geq 0 \\ -\sqrt{\ln(1-v)} & v \leq 0 \end{cases}$$

shows that full range may be satisfied even when $(h'(v))^{-1}$ is not bounded by a linear function.

[26]Previous existence results for matching models with imperfectly transferable utility and a continuum of types provided by Legros and Newman (2007b) and Kaneko and Wooders (1996) are not directly applicable because, in contrast to Legros and Newman (2007b), we avoid smoothness assumptions and, in contrast to Kaneko and Wooders (1996), our definition of a stable outcome given below insists on feasibility for all types. We also note that the powerful tools from the optimal transportation literature (c.f. Villani, 2009) are not available to us because these rely on quasilinearity. Section 6.1 of Nöldeke and Samuelson (2014) provides a more extensive discussion of related existence results.

Proposition 5 establishes that we can implement any increasing assignment, and satisfying feasibility is then a simple matter of choosing the appropriate increasing assignment to implement. Ensuring that we can satisfy participation constraints on both sides of the market requires an additional commonly-invoked assumption, requiring that matches are sufficiently productive. We then use Proposition 5 to ensure that we can implement a matching in which participation constraints are satisfied for buyers and bind for some particular buyer, and couple this with our productivity assumption to ensure that seller participation constraints are satisfied as well.

6.1 The Matching Model

As indicated in Section 2, we now interpret X as a set of buyers and Y as a set of sellers and the generating function ϕ and its inverse ψ as descriptions of the utility possibilities available to a matched pair of agents. Throughout the following, Assumption 1 is imposed. The sets X and Y are equipped with strictly positive Borel measures μ and ν, assumed to be non-atomic and satisfying $\mu(X) = \nu(Y)$. A *matching* is a measure-preserving bijection from X to Y which we may thus describe by a pair of assignments $(\boldsymbol{x}, \boldsymbol{y}) \in X^Y \times Y^X$ that are inverse to each other ($x = \boldsymbol{x}(y) \Longleftrightarrow y = \boldsymbol{y}(x)$ holds for all $x \in X$ and $y \in Y$) and measure-preserving.[27]

A matching together with a pair of utility profiles $(\boldsymbol{u}, \boldsymbol{v}) \in \boldsymbol{B}(X) \times \boldsymbol{B}(Y)$ is an *outcome*. An outcome $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$ is *feasible* if it satisfies the following equivalent conditions:

$$\begin{aligned} \boldsymbol{u}(x) &= \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x))) \quad \forall x \in X \\ \boldsymbol{v}(y) &= \psi(y, \boldsymbol{x}(y), \boldsymbol{u}(\boldsymbol{x}(y))) \quad \forall y \in Y, \end{aligned} \tag{44}$$

where the equivalence is due to the inverse relationship between $\boldsymbol{x}$ and $\boldsymbol{y}$ on the one hand and between ϕ and ψ on the other hand.

A feasible outcome is *individually rational* it satisfies

$$\boldsymbol{u} \geq \underline{\boldsymbol{u}} \text{ and } \boldsymbol{v} \geq \underline{\boldsymbol{v}}, \tag{45}$$

where $\underline{\boldsymbol{u}} \in \boldsymbol{CB}(X)$ and $\underline{\boldsymbol{v}} \in \boldsymbol{CB}(Y)$ are reservation utility profiles with the interpretation that $\underline{\boldsymbol{u}}(x)$ is the utility that buyer x can obtain when choosing her outside option of staying unmatched and $\underline{\boldsymbol{v}}(y)$ is the utility that seller y can obtain when choosing the outside option of staying unmatched. We impose the standard assumption that matches are productive in the sense that reservation utilities are feasible within any match. The interpretation is that matches may be pro-forma in the sense that both agents in a match could simply pursue their outside options.

[27]Note that the (standard) notion of an inverse relationship between $\boldsymbol{x}$ and $\boldsymbol{y}$ employed here is more restrictive than the one we considered when discussing the example in Section 3.5.

Assumption 4. *For all* $(x,y) \in X \times Y$ *the reservation utility profiles* $(\boldsymbol{u}, \boldsymbol{v})$ *satisfy the following equivalent conditions for all* $(x,y) \in X \times Y$*:*

$$\begin{aligned} \underline{\boldsymbol{u}}(x) &\leq \phi(x, y, \underline{\boldsymbol{v}}(y)) \\ \underline{\boldsymbol{v}}(y) &\leq \psi(y, x, \underline{\boldsymbol{u}}(x)). \end{aligned} \tag{46}$$

Provided that a measure preserving bijection $\boldsymbol{y}$ between X and Y exists, Assumption 4 ensures the existence of an individually rational outcome: simply assigning the reservation utility $\underline{\boldsymbol{v}}(y)$ to all sellers (that is, setting $\boldsymbol{v} = \underline{\boldsymbol{v}}$) and setting $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \underline{\boldsymbol{v}}(\boldsymbol{y}(x)))$ will do.

A feasible outcome is *stable* if it is individually rational and satisfies the equivalent pairwise stability conditions

$$\begin{aligned} \boldsymbol{u}(x) &\geq \phi(x, y, \boldsymbol{v}(y)) \ \ \forall x \in X \text{ and } y \in Y \\ \boldsymbol{v}(y) &\geq \psi(y, x, \boldsymbol{u}(x)) \ \ \forall x \in X \text{ and } y \in Y, \end{aligned} \tag{47}$$

where the equivalence is due to the inverse relationship between ϕ and ψ. Before proceeding we observe that combining the equalities in (44) with the inequalities in (47) shows that an individually rational outcome $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$ is stable if and only if $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ or, equivalently, $\boldsymbol{u}$ implements $(\boldsymbol{v}, \boldsymbol{x})$.[28] Hence, to establish the existence of a stable outcome it suffices to find a measure preserving bijection $\boldsymbol{y}$ and a pair of utility profiles $(\boldsymbol{u}, \boldsymbol{v})$ such that (i) $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ and (ii) the individual rationality condition (45) holds.

It is clear from the literature that establishing the existence of stable outcomes will require some additional structure. For example, we cannot find an increasing (as must be the case if the strict single-crossing condition of Assumption 3 holds) bijection between $X = [0, 2]$ and $Y = [0, 1] \cup [2, 3]$. Our response is to assume that X and Y are intervals.

We explain in Remarks 6 and 7 how the assumption that the measures μ and ν are atomless can be relaxed and how one could address the case in which there are different measures of buyers and sellers.

6.2 Existence Result

Suppose $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$ are intervals in $\mathbb{R}$ and let $m = \mu(X) = \nu(Y) > 0$. Then the distribution functions $F : X \to [0, m]$ and $G : Y \to [0, m]$ defined by

$$F(x) = \mu([\underline{x}, x]) \text{ and } G(y) = \nu([\underline{y}, y])$$

[28]The equivalence between the condition that $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ and that $\boldsymbol{u}$ implements $(\boldsymbol{v}, \boldsymbol{x})$ may seem puzzling at first glance as, in general, a profile $\boldsymbol{v}$ implementing a given pair $(\boldsymbol{u}, \boldsymbol{y})$ does not have to be implementable itself. What is special about the situation considered here is that $\boldsymbol{y}$ is a bijection, ensuring the equivalence between the conditions in (44).

are continuous (because μ and ν are non-atomic) and strictly increasing (because μ and ν are strictly positive) satisfying $F(\underline{x}) = G(\underline{y}) = 0$ and $F(\overline{x}) = G(\overline{y}) = m$. The assignment $\boldsymbol{y}^* \in Y^X$ defined by $\boldsymbol{y}^*(x) = G^{-1}(F(x))$ is then an increasing, measure-preserving bijection and hence, together with its inverse $\boldsymbol{x}^*$ describes a matching. This matching is *positive assortative*: higher buyers are matched with higher sellers. Provided that the single crossing condition from Assumption 3 holds, it is also stable in the sense that there exists utility profiles $(\boldsymbol{u}, \boldsymbol{v})$ such that $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{u}, \boldsymbol{v})$ is a stable outcome of the matching model. The proof is almost immediate from our previous results.

Proposition 8. *Let Assumptions 1, 3, and 4 hold and suppose X and Y are intervals with $\mu(X) = \nu(Y) > 0$. A stable outcome $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$ in which the matching is positive assortative exists.*

Proof. Let $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ be the positive assortative matching described before the statement of the proposition. As $\boldsymbol{y}^*$ is increasing, it is strongly implementable (Proposition 7). By Proposition 5 there exists $\boldsymbol{u} \in \boldsymbol{B}(X)$ satisfying $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ such that $(\boldsymbol{u}, \boldsymbol{y}^*)$ is implementable and satisfies $\boldsymbol{u}(\hat{x}) = \underline{\boldsymbol{u}}(\hat{x})$ for some $\hat{x} \in X$. Let $\boldsymbol{v} = \Psi\boldsymbol{u}$. By Corollary 3, $\boldsymbol{v}$ then implements $(\boldsymbol{u}, \boldsymbol{y}^*)$. It remains to verify that $\boldsymbol{v}$ satisfies the individual rationality constraint $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$ of the seller. Using Assumption 4 for the final inequality we have

$$\boldsymbol{v}(y) = \Psi\boldsymbol{u}(y) \geq \psi(y, \hat{x}, \boldsymbol{u}(\hat{x})) = \psi(y, \hat{x}, \underline{\boldsymbol{u}}(\hat{x})) \geq \underline{\boldsymbol{v}}(y),$$

finishing the proof. ■

Conversely, if the strict single-crossing property holds, then every stable outcome must be positive assortative.

Remark 6 (Atoms)**.** Proposition 8 continues to hold when the assumption that the measures μ and ν are non-atomic is dropped. We use this assumption to obtain the continuity of the distribution functions F and G defined on the intervals X and Y. In case F and/or G is not continuous Hellwig (2010) shows how the type spaces X and Y may be transformed (while preserving continuity of ϕ and the single crossing condition—which are the essential properties for our purposes) in such a way that the transformed types have continuous distribution function F' and G' on some compact intervals X' and Y', so that Proposition 8 is applicable. ■

Remark 7 (Monotonicity and Unbalanced Markets)**.** Legros and Newman (2007a) and Legros and Newman (2007b) assume that the reservation utility profiles $\underline{\boldsymbol{u}}$ and $\underline{\boldsymbol{v}}$ are identically equal to zero and that ϕ and ψ are increasing in their first two arguments, so that higher types of buyers and sellers are more productive. In this case assuming $\phi(\underline{x}, \underline{y}, 0) \geq 0$ is sufficient for Assumption 4 to hold and, further, Proposition 8 can be strengthened to ensure the existence of a stable outcome in which the participation constraint holds with equality for either buyer $\underline{x}$ or seller $\underline{y}$. Proposition 8 then continues to hold when the measures $\mu(X)$ and $\nu(Y)$ differ (and the definition of a stable outcome is extended in the obvious way to allow for agents to remain unmatched, with unmatched agents obtaining their reservation utility). Suppose for instance that $\mu(X) > \nu(Y)$ holds. Then the construction in the proof of Proposition 8 can be applied to the set $X' = [\underline{x}', \overline{x}]$, where $\underline{x}'$ is determined by the condition $\mu(X') = \nu(Y)$ to obtain a stable outcome for the model in which the types spaces are given by X' and Y and the condition $\boldsymbol{u}(\underline{x}') = 0$ holds. The single crossing condition and monotonicity then ensure that all buyers in the interval $[\underline{x}, \underline{x}')$ prefer staying unmatched and receiving their reservation utility of zero to any match acceptable for a seller, ensuring the stability of the resulting outcome. ■

Remark 8 (Multiple Stable Outcomes)**.** If (46) holds with strict inequality for all pairs (x, y), there will be multiple stable outcomes featuring positive assortative matching: The proof of Proposition 8 ensures that there exists such an outcome in which $\boldsymbol{u}(\hat{x}) = \underline{\boldsymbol{u}}(\hat{x})$ for some $\hat{x}$. We can reverse the roles of the buyer and seller in the proof of Proposition 8 to similarly ensure that there exists such an outcome in which $\boldsymbol{v}(\hat{y}) = \underline{\boldsymbol{v}}(\hat{y})$ for some $\hat{y}$. Moreover, these cannot be the same outcome. If so, we would have $(\boldsymbol{u}(\hat{x}), \boldsymbol{v}(\hat{y})) = (\underline{\boldsymbol{u}}(\hat{x}), \underline{\boldsymbol{v}}(\hat{y}))$, which from (44) contradicts strict inequality in (46) for the pair $(\hat{x}, \hat{y})$.

The existence of multiple stable outcomes raises the question whether something interesting can be said about the structure of the set of stable outcomes. The standard result in the matching literature (see Demange and Gale (1985, Section 4) or Roth and Sotomayor (1990, Chapter 9) for the case of a matching model with imperfectly transferable utility under consideration here, albeit with a finite number of agents) is that – using the dual order on either $\boldsymbol{B}(X)$ or $\boldsymbol{B}(Y)$ – the set of stable profiles $(\boldsymbol{u}, \boldsymbol{v})$ is a complete lattice. Strengthening the conditions under which Proposition 8 holds to include the strict single-crossing condition, this result immediately carries over to the case under consideration here. First, it follows from Lemma 8 and the fact that under the strict single crossing property every stable outcome exhibits the same (positive assortative) matching that the set of stable profiles is a lattice. Second, a lattice is complete if and only if it is compact in the interval topology (Birkhoff, 1995, p. 250, Theorem 20). The interval topology is weaker than our norm topology, and hence it suffices to show that the set of stable profiles is compact in $\boldsymbol{B}(X) \times \boldsymbol{B}(Y)$. Since that utility profiles in all stable profiles are implementable, closedness of the set of stable profiles in $\boldsymbol{B}(X) \times \boldsymbol{B}(Y)$ follows from an argument analogous to the one establishing Corollary 4. Invoking the Arzela-Ascoli theorem (Ok, 2007, p.264) and Corollary 5 it then suffices to show that the set of stable profiles is bounded to obtain the desired conclusion. We then note that the set of profiles $\boldsymbol{u}$ that are part of a stable outcome is bounded below by the participation constraint $\underline{\boldsymbol{u}}$ and is bounded above by $\Phi\underline{\boldsymbol{v}}$, with analogous bounds on the set of profiles $\boldsymbol{v}$. ■

7 Conclusion

We have introduced and studied a duality relationship that provides a characterization of implementable profiles and assignments in adverse-selection principal-agent models and two-sided matching models. This has allowed us to extend results previously developed for the quasilinear case, and to clarify the logic behind these results. We believe the techniques introduced here will find other applications.

Throughout our analysis we have eschewed smoothness assumptions, as these ought to play no role in the kind of characterization and existence results pursued here. However, much of the power of convex analysis stems from the inherent smoothness properties of convex functions, and many of the more useful implications of generalized convex analysis for the quasilinear case (e.g., the familiar integral representation of the agent's utility profile) require smoothness conditions. Adding such conditions to our Assumption 1 opens the possibility to investigate questions that go beyond those addressed in this paper. For instance, given our characterization of implementable assignments under the single crossing condition, we believe that (under suitable differentiability assumptions on the agents' utility function) the type-assignment approach developed in Nöldeke and Samuelson (2007) can be extended to the general case.

Appendix

A.1 Proof of Lemma 1

That ψ is strictly decreasing in its third argument for all $(y, x) \in Y \times X$ is immediate from (1) and the corresponding property of the generating function ϕ stated in Assumption 1. Because ϕ is defined on $X \times Y \times \overline{\mathbb{R}}$, we have $\psi(y, x, \overline{\mathbb{R}}) = \overline{\mathbb{R}}$ for all $(y, x) \in Y \times X$. Except for a permutation of the arguments, the epigraph (hypograph) of ϕ coincides with the hypograph (epigraph) of ψ. As a function into the extended real numbers is continuous if and only if its epigraph and hypograph are closed (Aliprantis and Border, 2006, Corollary 2.60, p. 52), it follows that the continuity of ϕ is equivalent to the continuity of ψ.

A.2 Proof of Lemma 2

Let $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ be implementable. (The proof for the dual case of implementable $(\boldsymbol{v}, \boldsymbol{x}) \in \mathbb{R}^Y \times X^Y$ is analogous.)

Fix $\boldsymbol{v} \in \overline{\mathbb{R}}^Y$ implementing $(\boldsymbol{u}, \boldsymbol{y})$. Then $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$ holds for all $x \in X$. Because $\boldsymbol{u}(x)$ is finite for all $x \in X$ and ϕ is strictly decreasing in its third argument, it follows that the equation

$$\boldsymbol{t}(x) = \boldsymbol{v}(\boldsymbol{y}(x)) \quad \forall x \in X$$

defines a real-valued function $\boldsymbol{t} : X \to \mathbb{R}$. Because $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ we have (using the inverse relationship between ϕ and ψ for the equivalence)

$$\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x)) \iff \boldsymbol{t}(x) = \psi(\boldsymbol{y}(x), x, \boldsymbol{u}(x)) \quad \forall x \in X \tag{48}$$

as well as the incentive compatibility conditions

$$\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x)) \geq \phi(x, \boldsymbol{y}(\tilde{x}), \boldsymbol{t}(\tilde{x})) \tag{49}$$

$$\boldsymbol{u}(\tilde{x}) = \phi(\tilde{x}, \boldsymbol{y}(\tilde{x}), \boldsymbol{t}(\tilde{x})) \geq \phi(\tilde{x}, \boldsymbol{y}(x), \boldsymbol{t}(x)) \tag{50}$$

for all $x, \tilde{x} \in X$.

First, we establish the existence of $\underline{\boldsymbol{u}}, \overline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ such that $\overline{\boldsymbol{u}} \geq \boldsymbol{u} \geq \underline{\boldsymbol{u}}$ holds, showing $\boldsymbol{u} \in \boldsymbol{B}(X)$. Fix any $\tilde{x} \in X$. From (49) we have $\boldsymbol{u}(x) \geq \phi(x, \boldsymbol{y}(\tilde{x}), \boldsymbol{t}(\tilde{x}))$ for all $x \in X$. Hence, defining $\underline{\boldsymbol{u}}$ by $\underline{\boldsymbol{u}}(x) = \phi(x, \boldsymbol{y}(\tilde{x}), \boldsymbol{t}(\tilde{x}))$, we have $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$. Further, as ϕ is continuous in x, X is compact, and $\boldsymbol{t}(\tilde{x})$ is finite, we have $\underline{\boldsymbol{u}} \in \boldsymbol{B}(X)$. Next, to construct $\overline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ satisfying $\overline{\boldsymbol{u}} \geq \boldsymbol{u}$, let $\underline{t} = \min_{y \in Y} \psi(y, \tilde{x}, \boldsymbol{u}(\tilde{x}))$. Because ψ is continuous in y and Y is compact, we have $\underline{t} \in \mathbb{R}$. Using the inverse relationship between ϕ and ψ and the strict monotonicity of the generating function in its third argument, the equivalence in (48) and the inequality in (50) imply $\boldsymbol{t}(x) \geq \psi(\boldsymbol{y}(x), \tilde{x}, \boldsymbol{u}(\tilde{x}))$. Hence, we have

$$\boldsymbol{t}(x) = \psi(\boldsymbol{y}(x), x, \boldsymbol{u}(x)) \geq \psi(\boldsymbol{y}(x), \tilde{x}, \boldsymbol{u}(\tilde{x})) \geq \underline{t} \tag{51}$$

for all $x \in X$. From the inequality $\boldsymbol{t}(x) \geq \underline{t}$ and the equality in (49) we have $\boldsymbol{u}(x) \leq \phi(x, \boldsymbol{y}(x), \underline{t})$ for all $x \in X$. Hence, defining $\overline{\boldsymbol{u}}$ by setting $\overline{\boldsymbol{u}}(x) = \phi(x, \boldsymbol{y}(x), \underline{t})$, we have $\overline{\boldsymbol{u}} \geq \boldsymbol{u}$. Further, as ϕ is continuous in (x, y), $X \times Y$ is compact, and $\underline{t}$ is finite, we have $\overline{\boldsymbol{u}} \in \boldsymbol{B}(X)$.

Second, we show there exists $\tilde{\boldsymbol{v}} \in \boldsymbol{B}(Y)$ implementing $(\boldsymbol{u}, \boldsymbol{y})$. Because $\boldsymbol{u} \in \boldsymbol{B}(X)$ holds, there exists $\underline{u}, \overline{u} \in \mathbb{R}$ such that $\overline{u} \geq \boldsymbol{u}(x) \geq \underline{u}$ holds for all $x \in X$. By compactness of X and continuity of ψ, setting $\overline{\boldsymbol{v}}(y) = \max_{x \in X} \psi(y, x, \underline{u})$ and $\underline{\boldsymbol{v}}(y) = \min_{x \in X} \psi(y, x, \overline{u})$ for all $y \in Y$ then defines real-valued profiles $\overline{\boldsymbol{v}} \in \mathbb{R}^Y$ and $\underline{\boldsymbol{v}} \in \mathbb{R}^Y$. Further, using compactness of Y, Berge's maximum theorem (Ok, 2007, p. 306) implies that these profiles are continuous and thus bounded. Because $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$, we have $\underline{\boldsymbol{v}}(y) \leq \boldsymbol{v}(y) \leq \overline{\boldsymbol{v}}(y)$ for all y in the range of $\boldsymbol{y}$. It follows that the profile $\tilde{\boldsymbol{v}}$ defined by $\tilde{\boldsymbol{v}}(y) = \boldsymbol{v}(y)$ for y in the range of $\boldsymbol{y}$ and $\tilde{\boldsymbol{v}}(y) = \overline{\boldsymbol{v}}(y)$ otherwise satisfies $\tilde{\boldsymbol{v}} \in \boldsymbol{B}(Y)$. Further, by (49) and the definition of $\overline{\boldsymbol{v}}$ the profile $\tilde{\boldsymbol{v}}$ implements $(\boldsymbol{u}, \boldsymbol{y})$.

A.3 Proof of Lemma 3

Suppose that $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is lower semicontinuous. Then the epigraph $E_{\boldsymbol{v}} = \{(y, v) \in Y \times \mathbb{R} \mid v \geq v(y)\}$ of $\boldsymbol{v}$ is closed. Because $\boldsymbol{v}$ is bounded, there exist $\underline{v} < \overline{v} \in \mathbb{R}$ with $\underline{v} \leq \boldsymbol{v}(y) \leq \overline{v}$ for all $y \in Y$. The set $Z_{\boldsymbol{v}} = E_{\boldsymbol{v}} \cap [Y \times \{[\underline{v}, \overline{v}]\}$ is then a closed subset of the compact set $Y \times [\underline{v}, \overline{v}]$ and hence is compact. The generating function ϕ is continuous (Assumption 1), and so a solution to the problem

$$\max_{(y,v)\in Z_{\boldsymbol{v}}} \phi(x, y, v) \tag{52}$$

exists for all $x \in X$ by Weierstrass' extreme value theorem. The graph of $\boldsymbol{v}$ is contained in $Z_{\boldsymbol{v}}$ and (because ϕ is strictly decreasing in its third argument) any solution to (52) lies on the graph of $\boldsymbol{v}$. Hence, for all $x \in X$, we have

$$\max_{(y,v)\in Z_{\boldsymbol{v}}} \phi(x, y, v) \geq \phi(x, y, \boldsymbol{v}(y)) \quad \forall y \in Y.$$

This ensures that the suprema in the definition of $\Phi\boldsymbol{v}$ are obtained, so that $\boldsymbol{v}$ implements $\boldsymbol{u} = \Phi\boldsymbol{v}$.

Using Lemma 1 to obtain the requisite properties of the inverse generating function ψ, the argument showing that a lower semicontinuous $\boldsymbol{u} \in \boldsymbol{B}(X)$ implements $\boldsymbol{v} = \Psi\boldsymbol{u}$ is analogous.

A.4 Proof of Lemma 4

We prove the first equality in (17) and the first inclusion in (18); using Lemma 1 to obtain the requisite properties of the inverse generating function ψ, the argument for the other cases is dual.

Lemma 2 establishes $\boldsymbol{I}(X) \subset \Phi\boldsymbol{B}(Y)$. Hence, it remains to show that for $\boldsymbol{v} \in \boldsymbol{B}(Y)$ the profile $\boldsymbol{u} = \Phi\boldsymbol{v}$ is implementable and continuous.

Suppose $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is lower semicontinuous. Then the implementability of $\boldsymbol{u} = \Phi\boldsymbol{v}$ follows from Lemma 3. Further, defining the set $Z_{\boldsymbol{v}}$ as in the proof of Lemma 3, we have $\boldsymbol{u}(x) = \max_{(y,v)\in Z_{\boldsymbol{v}}} \phi(x, y, v)$ for all $x \in X$. As ϕ is continuous and $Z_{\boldsymbol{v}}$ is compact, continuity of $\boldsymbol{u}$ then follows from Berge's maximum theorem (Ok, 2007, p. 306).

If $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is not lower semicontinuous, consider the lower semicontinuous hull $\bar{\boldsymbol{v}}$ of $\boldsymbol{v}$, i.e., the greatest element of the family of lower semicontinuous functions from Y to the extended real numbers $\bar{\mathbb{R}}$ majorized by $\boldsymbol{v}$. (The existence of $\bar{\boldsymbol{v}}$ is assured, cf. Penot (2013, Proposition 1.21).) As $\boldsymbol{v}$ is bounded, so is $\bar{\boldsymbol{v}}$, i.e., we have $\bar{\boldsymbol{v}} \in \boldsymbol{B}(Y)$. From Lemma 3 $\bar{\boldsymbol{v}}$ implements $\bar{\boldsymbol{u}} = \Phi\bar{\boldsymbol{v}}$. From the argument in the previous paragraph $\bar{\boldsymbol{u}}$ is continuous. It remains to show that $\bar{\boldsymbol{u}} = \boldsymbol{u}$ holds. Because the epigraph of $\bar{\boldsymbol{v}}$ is the closure of the epigraph of $\boldsymbol{v}$ (Penot, 2013, Proposition 1.21), we have

$$\sup_{(y,v)\in Z_{\boldsymbol{v}}} \phi(x, y, v) = \max_{(y,v)\in Z_{\bar{\boldsymbol{v}}}} \phi(x, y, v)$$

and thus $\sup_{y\in Y} \phi(x, y, \boldsymbol{v}(y)) = \max_{y\in Y} \phi(x, y, \bar{\boldsymbol{v}}(y))$ for all $x \in X$, which is the desired result.

A.5 Proof of Lemma 5

[5.1] We prove the continuity of $\Psi : \boldsymbol{B}(X) \to \boldsymbol{B}(Y)$. The argument for the continuity of $\Phi : \boldsymbol{B}(Y) \to \boldsymbol{B}(X)$ is analogous.

Fix $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\epsilon > 0$. We have to establish that there exists $\delta > 0$ such that

$$\|\tilde{\boldsymbol{u}} - \boldsymbol{u}\| < \delta \Longrightarrow \|\Psi\tilde{\boldsymbol{u}} - \Psi\boldsymbol{u}\| < \epsilon.$$

Let (the following expressions are well-defined because $\boldsymbol{u}$ is bounded) $\bar{z} = \sup_{x\in X} \boldsymbol{u}(x) + 1$, $\underline{z} = \inf_{x\in X} \boldsymbol{u}(x) - 1$, and $Z = [\underline{z}, \bar{z}] \subset \mathbb{R}$. For every $\delta < 1$ and $x \in X$, we then have

$$\|\tilde{\boldsymbol{u}} - \boldsymbol{u}\| < \delta \Longrightarrow \tilde{\boldsymbol{u}}(x) \in Z.$$

As ψ is continuous (Lemma 1), it is uniformly continuous on the compact set $X \times Y \times Z$. Hence, there exists $0 < \delta < 1$ such that

$$\|\tilde{\boldsymbol{u}} - \boldsymbol{u}\| < \delta \Longrightarrow \|\psi(y, x, \tilde{\boldsymbol{u}}(x)) - \psi(y, x, \boldsymbol{u}(x))\| < \epsilon$$

for all $x \in X$ and $y \in Y$. We also have

$$\begin{gathered}\|\psi(y, x, \tilde{\boldsymbol{u}}(x)) - \psi(y, x, \boldsymbol{u}(x))\| < \epsilon \;\text{ for all } x \in X \text{ and } y \in Y \Longrightarrow \\ \|\sup_{x\in X} \psi(y, x, \tilde{\boldsymbol{u}}(x)) - \sup_{x\in X} \psi(y, x, \boldsymbol{u}(x))\| \leq \epsilon \;\text{ for all } y \in Y.\end{gathered}$$

We thus have

$$\|\tilde{\boldsymbol{u}} - \boldsymbol{u}\| < \delta \Longrightarrow \|\Psi\tilde{\boldsymbol{u}}(y) - \Psi\boldsymbol{u}(y))\| < \epsilon \;\text{ for all } y \in Y,$$

proving $\|\Psi\tilde{\boldsymbol{u}} - \Psi\boldsymbol{u}\| < \epsilon$, as desired.

[5.2] Suppose $\mathcal{V} \subset \boldsymbol{B}(Y)$ is bounded, ensuring the existence of a compact interval $Z \subset \mathbb{R}$ such that $\boldsymbol{v}(Y) \subset Z$ holds for all $\boldsymbol{v} \in \mathcal{V}$. We then have $\Phi\boldsymbol{v}(x) \in [\min_{(x,y,v)\in X\times Y\times Z} \phi(x, y, v), \max_{(x,y,v)\in X\times Y\times Z} \phi(x, y, v)]$ for all $x \in X$ and $\boldsymbol{v} \in \mathcal{V}$, ensuring that $\Phi\mathcal{V} \subset \boldsymbol{B}(X)$ is bounded. An analogous argument yields that $\Psi\mathcal{U} \subset \boldsymbol{B}(Y)$ is bounded whenever $\mathcal{U} \subset \boldsymbol{B}(X)$ is bounded.

A.6 Proof of Lemma 6

Fix any $\boldsymbol{v} \in \boldsymbol{I}(Y)$. Let $(\boldsymbol{v}_n)$ be a sequence in $\boldsymbol{I}(Y)$ converging to $\boldsymbol{v}$. As the sequence $(\boldsymbol{v}_n)$ is bounded, Assumption 2 ensures we can apply Fatou's Lemma (Shiryaev, 1996, p. 187) to obtain

$$\limsup_{n\to\infty} \hat{\Pi}(\boldsymbol{v}_n) \leq \int_{x\in X} \limsup_{n\to\infty} \hat{\pi}(x, \boldsymbol{v}_n) d\mu(x).$$

Provided that the function $\hat{\pi}$ is upper semicontinuous in $\boldsymbol{v}$ we also have

$$\limsup_{n\to\infty} \hat{\pi}(x, \boldsymbol{v}_n) \leq \hat{\pi}(x, \boldsymbol{v}),$$

yielding $\limsup_{n\to\infty} \hat{\Pi}(\boldsymbol{v}_n) \leq \hat{\Pi}(\boldsymbol{v})$, which is the upper semicontinuity of $\hat{\Pi}$. It thus suffices to establish the upper semicontinuity of $\hat{\pi}$ in $\boldsymbol{v}$.

For any $x \in X$ define $w_x : Y \times \boldsymbol{I}(Y) \to \mathbb{R}$ by $w_x(y, \boldsymbol{v}) = \pi(x, y, \boldsymbol{v}(y))$ and $z_x : Y \times \boldsymbol{I}(Y) \to \mathbb{R}$ by $z_x(y, \boldsymbol{v}) = \phi(x, y, \boldsymbol{v}(y))$. We then have $\hat{\pi}(x, \boldsymbol{v}) = \max_{y \in \boldsymbol{Y_v}(x)} w_x(y, \boldsymbol{v})$ and $\boldsymbol{Y_v}(x) = \operatorname{argmax}_{y\in Y} z_x(y, \boldsymbol{v})$. As π and ϕ have been assumed to be continuous and $\boldsymbol{v}$ is continuous by Lemma 4.2, the functions w_x and z_x are jointly continuous in y and $\boldsymbol{v}$. Using continuity of z_x, Berge's maximum theorem (Ok, 2007, p. 306) implies that $\boldsymbol{Y_v}(x)$ is upper hemicontinuous in $\boldsymbol{v}$. As $\boldsymbol{Y_v}(x)$ is also non-empty and compact for all $x \in X$ and $\boldsymbol{v} \in \boldsymbol{I}(Y)$, the continuity of w_x then implies upper semicontinuity of $\hat{\pi}$ in $\boldsymbol{v}$ (Aliprantis and Border, 2006, Lemma 17.30).

A.7 Proof of Lemma 7

The first step of the proof is to show that if that if the entire tariff $\boldsymbol{v}(y)$ lies below some bound, then the principal's payoff falls short of $\Pi(\Psi\underline{\boldsymbol{u}})$, ensuring that such a tariff cannot be optimal.

Let $c < \hat{\Pi}(\Psi\underline{\boldsymbol{u}})$. By an argument analogous to the one proving Lemma 1, Assumption 2 ensures that the implicit equation $\pi(x, y, \rho(x, y)) = c$ defines a continuous function $\rho : X \times Y \to \mathbb{R}$. Let $\underline{\rho} = \min_{(x,y)\in X\times Y} \rho(x, y)$. For any $\boldsymbol{v} \in \boldsymbol{I}(Y)$ satisfying $\boldsymbol{v}(y) \leq \underline{\rho}$ for all $y \in Y$ we then have $\pi(x, y, \boldsymbol{v}(y)) \leq c < \hat{\Pi}(\Psi\underline{\boldsymbol{u}})$ and thus $\hat{\Pi}(\boldsymbol{v}) < \hat{\Pi}(\psi\underline{\boldsymbol{u}})$.

The second step of the proof shows that there exists a tariff $\underline{\boldsymbol{v}} \in \boldsymbol{I}(Y)$ such that if $\boldsymbol{v} \in \boldsymbol{I}(Y)$ fails $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$, then $\boldsymbol{v}(y) \leq \underline{\rho}$ for all $y \in Y$. This establishes $\underline{\boldsymbol{v}}$ as the required lower bound on the tariffs the principal need consider. Intuitively, the idea is that if $\boldsymbol{v}(\hat{y}) < \underline{\boldsymbol{v}}(\hat{y})$ for some $\hat{y}$, then the opportunity for agents to choose $\hat{y}$ imposes an upper bound on the tariff that an implementable profile $\boldsymbol{v}$ can attach to any other decision y, giving $\boldsymbol{v}(y) \leq \underline{\rho}$ for all $y \in Y$.

Define the function $h_n : X \times Y \times Y \to \mathbb{R}$ by

$$h_n(x, y, y') = \max\{\phi(x, y, \underline{\rho}) - \phi(x, y', \underline{\rho} - n), -1\}$$

Then $\{h_n\}_{,=0}^{\infty}$ is a decreasing sequence of continuous functions approaching a continuous limit (the constant function -1) on the compact set $X \times Y \times Y$, and hence (by Dini's theorem (Ok, 2007, p. 255)) converges uniformly to this limit. We can thus select a value n with

$$\phi(x, y, \underline{\rho}) < \phi(x, y', \underline{\rho} - n) \tag{53}$$

for all $(x, y, y') \in X \times Y \times Y$. Now let $\tilde{\boldsymbol{v}}$ be the constant function equal to $\underline{\rho} - n$ and let $\underline{\boldsymbol{v}} = \Psi\Phi(\tilde{\boldsymbol{v}})$. Then from Lemma 4.1 we have $\underline{\boldsymbol{v}} \in \boldsymbol{I}(Y)$ and from the cancellation property in Corollary 1.1 we have $\underline{\boldsymbol{v}}(y) \leq \underline{\rho} - n$ for all $y \in Y$.

Let $\boldsymbol{v} \in \boldsymbol{I}(Y)$ be such that there exists $\hat{y}$ with $\boldsymbol{v}(\hat{y}) < \underline{\boldsymbol{v}}(\hat{y})$, and let $\boldsymbol{u} = \Phi(\boldsymbol{v})$.

Then we have

$$
\begin{aligned}
\boldsymbol{v}(y) &= \max_{x\in X} \psi(y, x, \boldsymbol{u}(x)) \\
&\leq \max_{x\in X} \psi(y, x, \phi(x, \hat{y}, \boldsymbol{v}(\hat{y})) \\
&\leq \max_{x\in X} \psi(y, x, \phi(x, \hat{y}, \underline{\rho} - n)) \\
&\leq \max_{x\in X} \psi(y, x, \phi(x, y, \underline{\rho})) \\
&= \underline{\rho},
\end{aligned}
$$

where the first equality follows from the fact that $\boldsymbol{u}$ implements $\boldsymbol{v}$ (Proposition 3), the subsequent inequality from the fact that ψ is decreasing in its third argument and $\boldsymbol{u}(x) \geq \phi(x, \hat{y}, \boldsymbol{v}(\hat{y}))$, the next inequality from the fact that $\boldsymbol{v}(\hat{y}) < \underline{\boldsymbol{v}}(\hat{y}) \leq \underline{\rho} - n$ and the fact that ϕ and ψ are decreasing in their third arguments, the next inequality from the fact that ψ is decreasing in its third argument, and the final equality from the inverse relationship (2). This gives the required result.

A.8 Proof of Lemma 8

Let $\boldsymbol{v}_1$ implement $(\boldsymbol{u}_1, \boldsymbol{y})$ and $\boldsymbol{v}_2$ implement $(\boldsymbol{u}_2, \boldsymbol{y})$. We show that $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$ implements $(\boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{y})$. An analogous argument applies to the other three cases appearing in the statement of the lemma.

To simplify notation let

$$\bar{\boldsymbol{u}}(x) = \max\{\boldsymbol{u}_1(x), \boldsymbol{u}_2(x)\} \text{ and } \underline{\boldsymbol{v}}(y) = \min\{\boldsymbol{v}_1(y), \boldsymbol{v}_2(y)\}.$$

so that $\boldsymbol{u}_1 \vee \boldsymbol{u}_2 = \bar{\boldsymbol{u}}$ and $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2 = \underline{\boldsymbol{v}}$.

For all $x \in X$, we have

$$
\begin{aligned}
\bar{\boldsymbol{u}}(x) &= \max\{\phi(x, \boldsymbol{y}(x), \boldsymbol{v}_1(\boldsymbol{y}(x))), \phi(x, \boldsymbol{y}(x), \boldsymbol{v}_2(\boldsymbol{y}(x)))\} \\
&= \phi(x, \boldsymbol{y}(x), \underline{\boldsymbol{v}}(x)),
\end{aligned}
$$

where the first equality holds because $\boldsymbol{v}_1$ implements $\boldsymbol{u}_1$, $\boldsymbol{v}_2$ implements $\boldsymbol{u}_2$, and both $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ implement $\boldsymbol{y}$. The second equality holds because ϕ is decreasing in its third argument. Further, for all $(x, y) \in X \times Y$, we have

$$
\begin{aligned}
\bar{\boldsymbol{u}}(x) &\geq \max\{\phi(x, y, \boldsymbol{v}_1(y)), \phi(x, y, \boldsymbol{v}_2(y))\} \\
&= \phi(x, y, \underline{\boldsymbol{v}}(y)),
\end{aligned}
$$

where the inequality follows from the fact that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ implement $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ and the equality holds because ϕ is decreasing in its third argument. We thus have

$$\bar{\boldsymbol{u}}(x) = \phi(x, \boldsymbol{y}(x), \underline{\boldsymbol{v}}(\boldsymbol{y}(x))) = \max_{y\in Y} \phi(x, y, \underline{\boldsymbol{v}}(y))$$

for all $x \in X$, which is the statement that $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$ implements $(\boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{y})$.

A.9 Proof of Proposition 5

Let $\boldsymbol{y} \in Y^X$ be strongly implementable and $\underline{\boldsymbol{u}} \in \boldsymbol{B}(x)$ continuous. For $x \in X$ let $\mathcal{U}_x$ be the set of $\boldsymbol{u} \in \boldsymbol{B}(X)$ satisfying $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$ for which there exists $\boldsymbol{v} \in \boldsymbol{B}(Y)$ implementing $(\boldsymbol{u}, \boldsymbol{y})$. By hypothesis, the set $\mathcal{U}_x$ is non-empty for all $x \in X$.

In preparation for the following main argument, let $\mathcal{V}_x = \Psi\mathcal{U}_x$. We show that the set $\mathcal{V} = \cup_{x\in X}\mathcal{V}_x$ is bounded below and the set $\mathcal{U} = \cup_{x\in X}\mathcal{U}_x$ is bounded above. Let $\boldsymbol{v} \in \mathcal{V}$. Then there exists $x \in X$ and $\boldsymbol{u} \in \mathcal{U}_x$ implementing $\boldsymbol{v}$, yielding

$$\boldsymbol{v}(y) \geq \psi(y, x, \underline{\boldsymbol{u}}(x)) \geq \min_{\hat{x}\in X} \psi(y, \hat{x}, \underline{\boldsymbol{u}}(\hat{x})) \text{ for all } y \in Y,$$

where the minimum exists because ψ and $\underline{\boldsymbol{u}}$ are continuous (the former from Lemma 1 and the latter by assumption) and X is compact. Defining $\boldsymbol{v}_b \in \boldsymbol{B}(Y)$ by setting $\boldsymbol{v}_b(y) = \min_{x\in X} \psi(y, x, \underline{\boldsymbol{u}}(x))$, we then have $\boldsymbol{v} \geq \boldsymbol{v}_b$ for all $\boldsymbol{v} \in \mathcal{V}$, providing the lower bound on $\mathcal{V}$. Using (19) we have $\mathcal{U} = \Phi\mathcal{V}$. Hence, we may apply the order reversal property from Corollary 1.2 to infer that $\boldsymbol{u} \leq \Phi\boldsymbol{v}_b$ holds for all $\boldsymbol{u} \in \mathcal{U}$, showing that $\mathcal{U}$ is bounded above.

Pick an arbitrary $x_0 \in X$ and an implementable pair $(\boldsymbol{u}_0, \boldsymbol{v}_0) \in \mathcal{U}_{x_0} \times \mathcal{V}_{x_0}$. Define a sequence $(x_n, \boldsymbol{u}_n, \boldsymbol{v}_n)_{n=1}^{\infty}$ with $(x_n, \boldsymbol{u}_n, \boldsymbol{v}_n) \in X \times \mathcal{U}_{x_n} \times \mathcal{V}_{x_n}$ by the following recursion: given $(x_{n-1}, \boldsymbol{u}_{n-1}, \boldsymbol{v}_{n-1}) \in X \times \mathcal{U}_{x_{n-1}} \times \mathcal{V}_{x_{n-1}}$ let $x_n \in \arg\min_{x\in X} [\boldsymbol{u}_{n-1}(x) - \underline{\boldsymbol{u}}(x)]$. As both $\boldsymbol{u}_{n-1}(x)$ and $\underline{\boldsymbol{u}}(x)$ are continuous (the former by Lemma 4.2 and the latter by assumption), x_n is well-defined. Pick any $\tilde{\boldsymbol{u}}_n \in \mathcal{U}_{x_n}$ and $\tilde{\boldsymbol{v}}_n$ implementing $(\tilde{\boldsymbol{u}}_n, \boldsymbol{y})$. We then have $(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{v}}_n) \in \mathcal{U}_{x_n} \times \mathcal{V}_{x_n}$. Let $\boldsymbol{u}_n = \tilde{\boldsymbol{u}}_n \vee \boldsymbol{u}_{n-1}$ and $\boldsymbol{v}'_n = \tilde{\boldsymbol{v}}_n \wedge \boldsymbol{v}_{n-1}$. From Lemma 8 $\boldsymbol{v}'_n$ implements $(\boldsymbol{u}_n, \boldsymbol{y})$. Further, as $(x_{n-1}, \boldsymbol{u}_{n-1}) \in X \times \mathcal{U}_{x_{n-1}}$ holds, we have $\boldsymbol{u}_{n-1}(x_n) - \underline{\boldsymbol{u}}(x_n) \leq 0$, implying that $\boldsymbol{u}_n(x_n) = \max\{\boldsymbol{u}_{n-1}(x_n), \tilde{\boldsymbol{u}}_n(x_n)\} = \underline{\boldsymbol{u}}(x_n)$ holds. Consequently, we have $\boldsymbol{u}_n \in \mathcal{U}_{x_n}$. Letting $\boldsymbol{v}_n = \Psi u_n$ we then have $(\boldsymbol{u}_n, \boldsymbol{v}_n) \in \mathcal{U}_{x_n} \times \mathcal{V}_{x_n}$.[29]

By construction, the sequence $(\boldsymbol{u}_n)_{n=1}^{\infty}$ is an increasing sequence of continuous functions on the compact set X. Continuity is immediate from Lemma 4.2 because each $\boldsymbol{u}_n$ is implementable. The sequence is increasing because $\boldsymbol{u}_n = \tilde{\boldsymbol{u}}_n \vee \boldsymbol{u}_{n-1}$ implies $\boldsymbol{u}_n \geq \boldsymbol{u}_{n-1}$. Similarly, because each $\boldsymbol{v}_n = \Psi u_n$ is implementable and Ψ is order reversing, the sequence $(\boldsymbol{v}_n)_{n=1}^{\infty}$ is a decreasing sequence of continuous functions on the compact set Y. As $\mathcal{U}$ is bounded above and $\mathcal{V}$ is bounded below, it follows that $(\boldsymbol{u}_n, \boldsymbol{v}_n)_{n=1}^{\infty}$ converges pointwise to a limit that we denote by $(\boldsymbol{u}^*, \boldsymbol{v}^*)$. We now proceed in two steps.

First, $\boldsymbol{v}^*$ implements $(\boldsymbol{u}^*, \boldsymbol{y})$. This follows from noting that for all $n \geq 1$ and $(x, y) \in (X, Y)$ we have

$$\boldsymbol{u}_n(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}_n(\boldsymbol{y}(x))) \geq \phi(x, y, \boldsymbol{v}_n(y)).$$

Taking limits for fixed (x, y) we obtain

$$\boldsymbol{u}^*(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}^*(\boldsymbol{y}(x))) \geq \phi(x, y, \boldsymbol{v}^*(y))$$

[29] The argument used to prove Lemma 8 can be extended to show that we must have $v'_n = \Psi u_n$, so that this last step is actually superfluous.

for all $(x, y) \in (X, Y)$ which is the statement that $\boldsymbol{v}^*$ implements $(\boldsymbol{u}^*, \boldsymbol{y})$. Because $\boldsymbol{u}^*$ is implementable, it is continuous. It then follows from Dini's theorem (Ok, 2007, p. 255) that the sequence $(\boldsymbol{u}_n)_{n=1}^{\infty}$ converges to $\boldsymbol{u}^*$ uniformly.

Second, because X is compact, the sequence $(x_n)_{n=1}^{\infty}$ has a converging subsequence, denoted by x_{n_k}, with limit $x^* \in X$. As $(\boldsymbol{u}_n)_{n=1}^{\infty}$ is a sequence of continuous functions converging uniformly to $\boldsymbol{u}^*$ we have

$$\lim_{k\to\infty} \boldsymbol{u}_{n_k}(x_{n_k}) = \boldsymbol{u}^*(x^*) \tag{54}$$

$$\lim_{k\to\infty} \boldsymbol{u}_{n_k-1}(x_{n_k}) = \boldsymbol{u}^*(x^*). \tag{55}$$

As $\boldsymbol{u}_n(x_n) = \underline{\boldsymbol{u}}(x_n)$ holds for all n and $\underline{\boldsymbol{u}}$ is continuous, (54) implies

$$\boldsymbol{u}^*(x^*) = \underline{\boldsymbol{u}}(x^*). \tag{56}$$

By construction of the sequence $(x_n)_{n=1}^{\infty}$ we have

$$\boldsymbol{u}_{n-1}(x) - \underline{\boldsymbol{u}}(x) \geq \boldsymbol{u}_{n-1}(x_n) - \underline{\boldsymbol{u}}(x_n)$$

for all $x \in X$ and $n \geq 1$. Taking limits for the sequence n_k we thus obtain

$$\boldsymbol{u}^*(x) - \underline{\boldsymbol{u}}(x) \geq \boldsymbol{u}^*(x^*) - \underline{\boldsymbol{u}}(x^*)$$

for all $x \in X$, where we have used the continuity of $\underline{\boldsymbol{u}}$ and (55) to obtain the right side of the inequality. Taking account of (56) this implies

$$\boldsymbol{u}^*(x) \geq \underline{\boldsymbol{u}}(x) \tag{57}$$

for all $x \in X$. As $\boldsymbol{v}^*$ implements $(\boldsymbol{u}^*, \boldsymbol{y})$, (56)–(57) imply that $\boldsymbol{y}$ can be implemented with $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ and with $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$ for some x.

A.10 Proof of Lemma 10

The set $\boldsymbol{CB}_0(Y)$ is a non-empty, convex subset of the normed linear space $\boldsymbol{B}(Y)$. Hence, by a version of Schauder's fixed point (Granas and Dugundji, 2003, Theorem 3.2, p. 119) it suffices to show that the map $T : \boldsymbol{CB}_0(Y) \to \boldsymbol{CB}_0(Y)$ is continuous and is a compact map, meaning that $T\boldsymbol{CB}_0(Y)$ is contained in a compact subset of $\boldsymbol{CB}_0(Y)$ (Granas and Dugundji, 2003, Definition 1.1, page 112).

The mappings ρ and $\hat{\Phi}$ are clearly continuous and the implementation map Ψ is also continuous (Lemma 5). Hence, T is continuous. As $\boldsymbol{CB}_0(Y)$ is closed, it thus suffices to show that $T\boldsymbol{CB}_0(Y)$ is relatively compact to finish the proof.

Because relative compactness is preserved by continuous mappings and the mapping ρ is continuous, $T\boldsymbol{CB}_0(Y)$ is relatively compact if $\mathcal{V} = \Psi\hat{\Phi}\boldsymbol{CB}_0(Y)$ is relatively compact. By the Arzela-Ascoli theorem (Ok, 2007, p.264), $\mathcal{V}$ is relatively compact if it is bounded and equicontinuous. From Lemma 4.1, we have $\mathcal{V} \subset \boldsymbol{I}(Y)$.

Equicontinuity of $\mathcal{V}$ then follows from Corollary 5 provided that $\mathcal{V}$ is bounded. Hence, it remain to establish that $\mathcal{V}$ is bounded.

Let $\bar{\boldsymbol{v}} = \Psi\underline{\boldsymbol{u}}$. Because $\hat{\Phi}\boldsymbol{v} \geq \underline{\boldsymbol{u}}$ holds for all $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ and Ψ is order reversing (Corollary 1.2), we have $\bar{\boldsymbol{v}} \geq \Psi\hat{\Phi}\boldsymbol{v}$ for all $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$. Next, define $\underline{\boldsymbol{v}} \in \boldsymbol{B}(Y)$ by setting $\underline{\boldsymbol{v}}(y) = \psi(y, x_0, r_0)$ for all $y \in Y$. By construction of the set $\boldsymbol{CB}_0(Y)$, we have $\tilde{\Phi}\boldsymbol{v}(x_0) = r_0$ for all $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ and thus $\hat{\Phi}\boldsymbol{v}(x_0) = r_0$ for all $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$. Consequently, we have $\Psi\hat{\Phi}\boldsymbol{v}(y) \geq \psi(y, x_0, r_0) = \underline{\boldsymbol{v}}(y)$ for all $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ and thus

$$\bar{\boldsymbol{v}} \geq \Psi\hat{\Phi}\boldsymbol{v} \geq \underline{\boldsymbol{v}} \text{ for all } \boldsymbol{v} \in \boldsymbol{CB}_0(Y),$$

which is the boundedness of $\mathcal{V}$.

A.11 Proof of Proposition 6

By Lemma 10 there exists $\boldsymbol{v} \in \boldsymbol{CB}_0(Y)$ satisfying the fixed point condition $\boldsymbol{v} = T\boldsymbol{v} = \rho\Psi\hat{\Phi}\boldsymbol{v}$. We show that $\boldsymbol{v} = \Psi\tilde{\Phi}\boldsymbol{v}$. It then follows from Lemma 9 that $\boldsymbol{y}$ is implementable with initial condition (x_0, r_0).

We proceed in two steps. First, we address the effect of the function ρ. Let $\hat{\boldsymbol{u}} = \hat{\Phi}\boldsymbol{v}$ and $\hat{\boldsymbol{v}} = \Psi\hat{\boldsymbol{u}}$, so that $\hat{\boldsymbol{v}} = \Psi\hat{\Phi}\boldsymbol{v}$. By definition of the map $\hat{\Phi}$ and the conditions defining y_0 and v_0 we have $\hat{\boldsymbol{u}}(x_0) = r_0$ and thus

$$\hat{\boldsymbol{v}}(y_0) = \Psi\hat{\boldsymbol{u}}(y_0) = \sup_{x \in X} \psi(y_0, x, \hat{\boldsymbol{u}}(x)) \geq \psi(y_0, x_0, \hat{\boldsymbol{u}}(x_0)) = \psi(y_0, x_0, r_0) = v_0,$$

implying

$$\hat{\boldsymbol{v}}(y_0) \geq \boldsymbol{v}(y_0). \tag{58}$$

Using the order reversal property (Corollary 1.2), the inequality $\hat{\Phi}\boldsymbol{v} \geq \tilde{\Phi}\boldsymbol{v}$ implies $\Psi\hat{\Phi}\boldsymbol{v} \leq \Psi\tilde{\Phi}\boldsymbol{v}$. Using the assumption that there exists $\tilde{y} \in Y$ satisfying (38) it follows that there exists $\tilde{y} \in Y$ such that

$$\hat{\boldsymbol{v}}(\tilde{y}) \leq \boldsymbol{v}(\tilde{y}). \tag{59}$$

The definition of the mapping ρ implies that we have

$$\boldsymbol{v}(y_0) - \hat{\boldsymbol{v}}(y_0) = \boldsymbol{v}(y) - \hat{\boldsymbol{v}}(y) \tag{60}$$

for all $y \in Y$. From (58) the left side of (60) is nonpositive, whereas from (59) the right side of (60) is nonnegative for $y = \tilde{y}$. Hence, we have $\boldsymbol{v}(y) - \hat{\boldsymbol{v}}(y) = 0$ for all $y \in Y$, i.e., $\hat{\boldsymbol{v}} = \boldsymbol{v}$. This implies

$$\boldsymbol{v} = \Psi\hat{\Phi}\boldsymbol{v}. \tag{61}$$

Second, applying the implementation map Φ to both sides of (61) and using the cancellation rule from Corollary 1.1 yields the inequality $\hat{\Phi}\boldsymbol{v} \geq \Phi\boldsymbol{v}$. As

$$\Phi\boldsymbol{v}(x) \geq \phi(x, \boldsymbol{y}(x_0), \boldsymbol{v}(\boldsymbol{y}(x_0))) = \phi(x, y_0, v_0) > \phi(x, y_0, v_0) - \epsilon = \underline{\boldsymbol{u}}(x)$$

holds for all $x \in X$, the inequality $\hat{\Phi}\boldsymbol{v} \geq \Phi\boldsymbol{v}$ implies $\hat{\Phi}\boldsymbol{v} = \tilde{\Phi}\boldsymbol{v}$. We thus have $\boldsymbol{v} = \Psi\tilde{\Phi}\boldsymbol{v}$. Hence, it follows from Lemma 9 that $\boldsymbol{y}$ is implementable with initial condition (x_0, r_0).

A.12 Proof of Lemma 11

Let $\boldsymbol{v} \in \boldsymbol{B}(Y)$ be lower semicontinuous. By Lemma 3 this ensures that $\boldsymbol{Y_v}$ is nonempty valued. Assumption 3 implies that $w : X \times Y \to \mathbb{R}$ defined by $w(x, y) = \phi(x, y, \boldsymbol{v}(y))$ satisfies the single crossing property from Milgrom and Shannon (1994, p. 160), that is,

$$w(x_1, y_2) \geq w(x_1, y_1) \implies w(x_2, y_2) \geq w(x_2, y_1) \tag{62}$$

$$w(x_1, y_2) > w(x_1, y_1) \implies w(x_2, y_2) > w(x_2, y_1) \tag{63}$$

holds for all $x_1 < x_2 \in X$ and $y_1 < y_2 \in Y$.[30] This ensures that $\boldsymbol{Y_v}$ is nondecreasing in the strong set order (Milgrom and Shannon, 1994, Theorem 4), implying the existence of an increasing selection $\boldsymbol{y}$ from $\boldsymbol{Y_v}$ (Kukushkin, 2013, Theorem 2.7) which is then implemented by $\boldsymbol{v}$. Similarly, if ϕ satisfies the strict single crossing condition, then the second inequality in (62) is strict, implying that every selection from $\boldsymbol{Y_v}$ is increasing (Milgrom and Shannon, 1994, Theorem 4').

To obtain the dual statement, it clearly suffices to show that the inverse generating function ψ inherits the single crossing properties of the generating function ϕ. That is, we need to show that Assumption 3 implies

$$\psi(y_1, x_2, u_2) \geq \psi(y_1, x_1, u_1) \implies \psi(y_2, x_2, u_2) \geq \psi(y_2, x_1, u_1) \tag{64}$$

for all $y_1 < y_2 \in Y$, $x_1 < x_2 \in X$, and $u_1, u_2 \in \mathbb{R}$, with the second inequality holding strictly if ϕ satisfies the strict single crossing property.

Let $x_1 < x_2 \in X$, $y_1 < y_2 \in Y$, and let $u_1, u_2 \in \mathbb{R}$ be such that

$$\psi(y_1, x_2, u_2) \geq \psi(y_1, x_1, u_1) \tag{65}$$

holds. Let $v_2 = \psi(y_2, x_1, u_1)$ and let $v_1 = \psi(y_1, x_1, u_1)$. Using (1) we then have $\phi(x_1, y_2, v_2) = \phi(x_1, y_1, v_1)$. Applying (39) we obtain

$$\phi(x_2, y_2, v_2) \geq \phi(x_2, y_1, v_1). \tag{66}$$

Using the definition of v_1 and (65) we have $\psi(y_1, x_2, u_2) \geq v_1$. From (1) and the fact that ψ is strictly decreasing in its third argument, this implies $\phi(x_2, y_1, v_1) \geq u_2$, so that (66) yields $\phi(x_2, y_2, v_2) \geq u_2$. Because ϕ is strictly decreasing in its third argument, this implies $\psi(y_2, x_2, u_2) \geq v_2$. Using the definition of v_2 we thus obtain $\psi(y_2, x_2, u_2) \geq \psi(y_2, x_1, u_1)$, showing that (64) holds. The proof for the case in which ϕ satisfies the strict single crossing property is analogous.

[30] It is immediate from Assumption 3 that the first strict inequality in (63) implies that the second inequality must hold as a weak inequality. The strict inequality follows from the continuity and strict monotonicity of the generating function in its third argument imposed in Assumption 1.

A.13 Completion of the Proof of Proposition 7

If $\boldsymbol{y}$ is not continuous, consider the lower semicontinuous hull $\bar{\boldsymbol{u}}$ of $\tilde{\boldsymbol{u}}$. We have $\bar{\boldsymbol{u}}(x) = \liminf_{x_n \to x} \tilde{\boldsymbol{u}}(x_n)$ for all $x \in X$ (Penot, 2013, Proposition 1.21). Hence, for all $x \in X$ there exists a sequence x_n converging to x which (by taking a subsequence, if necessary) we may take to be monotonic (either increasing or decreasing), such that $\bar{\boldsymbol{u}}(x) = \lim_{x_n \to x} \phi(x_n, \boldsymbol{y}(x_n), \boldsymbol{v}(\boldsymbol{y}(x_n)))$. As $\boldsymbol{y}$ is increasing, the sequence $y_n = \boldsymbol{y}(x_n)$ is a bounded monotonic sequence and thus converges to some limit, which we denote by $\hat{\boldsymbol{y}}(x)$. The assignment $\hat{\boldsymbol{y}}$ is increasing.[31]

As $\boldsymbol{v}$ is continuous, $v_n = \boldsymbol{v}(y_n)$ converges to $\boldsymbol{v}(\hat{\boldsymbol{y}}(x))$. Hence, we have $\bar{\boldsymbol{u}}(x) = \phi(x, \hat{\boldsymbol{y}}(x), \boldsymbol{v}(\hat{\boldsymbol{y}}(x)))$ for all $x \in X$. Observing the equality $\Psi\tilde{\boldsymbol{u}} = \Psi\bar{\boldsymbol{u}}$ (cf. the proof of Lemma 4) we may now proceed exactly as in the proof for continuous $\boldsymbol{y}$ with $\tilde{\boldsymbol{u}}$ replaced by $\bar{\boldsymbol{u}}$: From Lemma 11, there exists an increasing $\boldsymbol{x} \in X^Y$ implemented by $\bar{\boldsymbol{u}}$, that is

$$\Psi\tilde{\Phi}\boldsymbol{v}(y) = \psi(y, \boldsymbol{x}(y), \bar{\boldsymbol{u}}(\boldsymbol{x}(y)))$$

holds for all $y \in Y$ and

$$\bar{\boldsymbol{u}}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{v}(\boldsymbol{y}(x)))$$

holds for all $x \in X$. We can again use Tarski's fixed point theorem to infer the existence of $(\tilde{x}, \tilde{y}) \in X \times Y$ satisfying

$$\begin{aligned} \Psi\tilde{\Phi}\boldsymbol{v}(\tilde{y}) &= \psi(\tilde{y}, \tilde{x}, \bar{\boldsymbol{u}}(\tilde{x})) \\ \bar{\boldsymbol{u}}(\tilde{x}) &= \phi(\tilde{x}, \tilde{y}, \boldsymbol{v}(\tilde{y})). \end{aligned}$$

Substituting the second of these equations into the first yields the desired equality $\Psi\tilde{\Phi}\boldsymbol{v}(\tilde{y}) = \boldsymbol{v}(\tilde{y})$.

References

Aliprantis, C., and K. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Verlag, third edn.

Balder, E. J. (1996): "On the Existence of Optimal Contract Mechanisms for Incomplete Information Principal-Agent Models," *Journal of Economic Theory*, 68(1), 133–148.

Bardsley, P. (2012): "Duality in Contracting," Discussion paper, University of Melbourne.

Birkhoff, G. (1995): *Lattice Theory*, vol. 25 of *Colloquium Publications*. American Mathematical Society, third edn.

[31] Let $x' < x''$. Then there are sequences x'_n and x''_n such that $\hat{\boldsymbol{y}}(x') = \lim_{x'_n \to x'} \boldsymbol{y}(x'_n)$ and $\hat{\boldsymbol{y}}(x'') = \lim_{x''_n \to x''} \boldsymbol{y}(x''_n)$. Because $x' < x''$, there exists N sufficiently large that by $x'_n < x''_n$, and hence $\boldsymbol{y}(x'_n) \le \boldsymbol{y}(x''_n)$, for all $n \ge N$, ensuring $\boldsymbol{y}(x') \le \boldsymbol{y}(x'')$.

Carbajal, J. C., and J. C. Ely (2013): "Mechanism Design without Revenue Equivalence," *Journal of Economic Theory*, 148(1), 104–133.

Carlier, G. (2001): "A General Existence Result for the Principal-Agent Problem with Adverse Selection," *Journal of Mathematical Economics*, 35(1), 129–150.

——— (2002): "Nonparametric Adverse Selection Problems," *Annals of Operations Research*, 114(1-4), 71–82.

Chiappori, P., R. McCann, and L. Nesheim (2010): "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," *Economic Theory*, 42(2), 317–354.

Davey, B., and H. Priestley (2002): *Introduction to Lattices and Order*. Cambridge University Press, 2 edn.

Demange, G., and D. Gale (1985): "The Strategy Structure of Two-Sided Matching Markets," *Econometrica*, 53(4), 873–888.

Ekeland, I. (2010a): "Notes on Optimal Transportation," *Economic Theory*, 42(2), 437–459.

Ekeland, I. (2010b): "Existence, Uniqueness and Efficiency of Equilibrium in Hedonic Markets with Multidimensional Types," *Economic Theory*, 42(2), 275–315.

Erné, M., J. Koslowski, A. Melton, and G. E. Strecker (1993): "A Primer on Galois Connections," *Annals of the New York Academy of Sciences*, 704(1), 103–125.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. MIT Press, Cambridge, Massachusetts.

Glosten, L. R. (1989): "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business*, 62, 211–235.

Goeree, J. K., and A. I. Kushnir (2013): "A Geometric Approach to Mechanism Design," Working Paper 56, Department of Economics, University of Zurich.

Granas, A., and J. Dugundji (2003): *Fixed Point Theory*. Springer, New York.

Guesnerie, R., and J. Laffont (1994): "Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *Journal of Public Economics*, 25, 329–369.

Hellwig, M. F. (1992): "Fully Revealing Outcomes in Signalling Models: An Example of Nonexistence when the Type Space is Unbounded," *Journal of Economic Theory*, 58(1), 93–104.

Hellwig, M. F. (2010): "Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach," *Econometrica*, 78(4), 1201–1237.

Heydenreich, B., R. Müller, M. Uetz, and R. Vohra (2009): "Characterization of Revenue Equivalence," *Econometrica*, 77(1), 307–316.

Jullien, B. (2000): "Participation Constraints in Adverse Selection Problems," *Journal of Economic Theory*, 93(1), 47.

Kadan, O., P. J. Reny, and J. Swinkels (2011): "Existence of Optimal Mechanisms in Principal-Agent Problems," Working Papers 2011-002, Becker Friedman Institute for Research In Economics, University of Chicago.

Kahn, C. M. (1993): "Existence and Characterization of Optimal Employment Contracts on a Continuous State Space," *Journal of Economic Theory*, 59(1), 122–144.

Kaneko, M., and M. Wooders (1996): "The Nonemptiness of the f-Core of a Game without Side Payments," *International Journal of Game Theory*, 25(2), 245–258.

Kukushkin, N. S. (2013): "Increasing Selections from Increasing Multifunctions," *Order*, 30(2), 541–555.

Legros, P., and A. Newman (2007a): "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities," *Econometrica*, 75(4), 1073–1102.

Legros, P., and A. Newman (2007b): "Assortative Matching in Two-sided Continuum Economies," Discussion paper.

Luenberger, D. G. (1992): "Benefit Functions and Duality," *Journal of Mathematical Economics*, 21(5), 461 – 481.

Mailath, G. J., and G. Nöldeke (2008): "Does Competitive Pricing Cause Market Breakdown under Extreme Adverse Selection?," *Journal of Economic Theory*, 140(1), 97–125.

Milgrom, P., and I. Segal (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583–601.

Milgrom, P., and C. Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

Mirrlees, J. A. (1986): "The Theory of Optimal Taxation," in *Handbook of Mathematical Economics*, ed. by K. J. Arrow, and M. Intriligator, vol. 3, chap. 24, pp. 1198–1249. Elsevier.

Monjardet, B. (1978): "An Axiomatic Theory of Tournament Aggregation," *Mathematics of Operations Research*, 3(4), pp. 334–351.

——— (2007): "Some Order Dualities in Logic, Games and Choices," *International Game Theory Review*, 9(1), 1–12.

Nöldeke, G., and L. Samuelson (2007): "Optimal Bunching without Optimal Control," *Journal of Economic Theory*, 134(1), 405–420.

Nöldeke, G., and L. Samuelson (2014): "Investment and Competitive Matching," Cowles Foundation Discussion Paper 1946, Yale University.

Ok, E. A. (2007): *Real Analysis with Economic Applications*. Princeton University Press.

Ore, O. (1944): "Galois Connexions," *Trans. Amer. Math. Soc*, 55(1944), 493–513.

Page, F. H. (1991): "Optimal Contract Mechanisms for Principal-Agent Problems with Moral Hazard and Adverse Selection," *Economic Theory*, 1(4), 323–338.

——— (1992): "Mechanism Design for General Screening Problems with Moral Hazard," *Economic Theory*, 2(2), 265–281.

——— (1997): "Optimal Deterministic Contracting Mechanisms for Principal-Agent Problems with Moral Hazard and Adverse Selection," *Review of Economic Design*, 3(1), 1–13.

Penot, J.-P. (2000): "What is Quasiconvex Analysis?," *Optimization*, 47(1-2), 35–110.

——— (2010): "Are Dualities Appropriate for Duality Theories in Optimization?," *Journal of Global Optimization*, 47(3), 503–525.

——— (2013): *Calculus without Derivatives*, vol. 266 of *Graduate Texts in Mathematics*. Springer, New York.

Rochet, J. (1985): "The Taxation Principle and Multi-time Hamilton-Jacobi Equations," *Journal of Mathematical Economics*, 14(2), 113–128.

Rochet, J. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-linear Context," *Journal of Mathematical Economics*, 16, 191–200.

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

Rockafellar, R. T., and R. J.-B. Wets (1998): *Variational Analysis*, vol. 317 of *Grundlehren der mathematischen Wissenschaften*. Springer.

Roth, A. E., and M. A. O. Sotomayor (1990): *Two-Sided Matching*. Cambridge University Press, Cambridge.

Shiryaev, A. N. (1996): *Probability*. Springer Verlag, second edn.

Singer, I. (1997): *Abstract Convex Analysis*. Wiley-Interscience.

Strausz, R. (2006): "Deterministic versus Stochastic Mechanisms in Principal-Agent Models," *Journal of Economic Theory*, 128(1), 306–314.

Villani, C. (2009): *Optimal Transport, Old and New*. Springer-Verlag, Berlin.

Vohra, R. V. (2011): *Mechanism Design: A Linear Programming Approach*, vol. 47. Cambridge University Press.

Weibull, J. (1989): "A Note on the Continuity of Incentive Schedules," *Journal of Public Economics*, 39, 239–243.