The Implementation Duality*

Georg Nöldeke[†] Larry Samuelson[‡]

October 29, 2017

Abstract: Conjugate duality relationships are pervasive in matching and implementation problems and provide much of the structure essential for characterizing stable matches and implementable allocations in models with quasilinear (or transferable) utility. In the absence of quasilinearity, a more abstract duality relationship, known as a Galois connection, takes the role of (generalized) conjugate duality. While much weaker, this duality relationship still induces substantial structure. We show that this structure can be used to extend existing results for, and gain new insights into, adverse-selection principal-agent problems and two-sided matching problems without quasilinearity.

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

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

*We thank Pierre-Andre Chiappori, Christopher Chambers, Vince Crawford, Christian Ewerhart, Alfred Galichon, Michael Greinecker, Samuel Häfner, Bruno Jullien, Igor Letina, Thomas Mariotti, Robert McCann, Sofia Moroni, Nick Netzer, Phil Reny, Frank Riedel, Jean-Charles Rochet, Christoph Schottmüller, Alex Teytelboym, the editor and five referees 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 and Nuffield College (Nöldeke) and All Souls College (Samuelson) at Oxford University. We are grateful to these institutions for their hospitality. Larry Samuelson thanks the National Science Foundation (SES-1459158) 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

Contents

The Implementation Duality

1 Introduction

Much of the theory of mechanism design with quasilinear utility can be developed from a linear programming perspective, with duality-based arguments taking center stage (Vohra, 2011). The fundamental duality of linear programming also plays a central role in the theory of matching models with quasilinear (transferable) utility, from the theory's inception in Shapley and Shubik (1972) to the more recent adoption of optimal transport methods (cf. Galichon, 2016) that are based on the Kantorovich duality for infinite dimensional linear programs (Villani, 2009). These are not disparate applications of linear programming. Rather, there is a deep connection between mechanism design and matching problems (in their guise as optimal transport problems) with quasilinear utility (Ekeland, 2010). Intuitively, this connection arises because stable outcomes in matching models with transferable utility are composed of optimal assignments (obtained as the solution to a primal linear programming problem) together with optimal utility profiles (obtained as the solution to the dual linear programming problem) that implement the optimal assignment in the sense of making it incentive compatible for all agents to choose their designated partners.

Models based on quasilinear utility, though analytically convenient, are ill-suited for mechanism design problems in which the stakes are sufficiently large to make income effects or risk aversion salient (e.g. Mirrlees, 1971; Stiglitz, 1977), and are also ill-suited for matching problems in which—either because of income effects or because of the structure of the underlying bilateral relationship (Legros and Newman, 2007; Chiappori and Salanié, 2016; Galichon, Kominers, and Weber, 2016)— utility is imperfectly transferable.

This paper studies implementation without invoking quasilinearity. In so doing, we lose access to the linear programming duality and hence cannot make use of it to characterize implementable assignments and their attended profiles. Our first step is to uncover a more abstract duality relationship, known as a Galois connection (Birkhoff, 1995, p. 124), that still induces substantial structure. In particular, it is well understood that the solution to the quasilinear optimal assignment problem gives rise to a second layer of duality—the optimal utility profiles are generalized conjugate duals of each other, and the optimal assignment is drawn from the argmax correspondence of the maximization problems inducing this duality (Galichon, 2016, Chapter 7). We exploit the properties of the Galois connection to show that these properties do not require quasilinear utility—implementability is in general characterized by utility profiles that are abstract conjugate duals of each other, and the assignment is drawn from the corresponding argmax correspondence. Moreover, these properties characterize pairwise stable outcomes in matching problems (with the addition of participation constraints then giving rise to a characterization of stable outcomes).

The first part of the paper, consisting of Sections 2 and 3, provides the theoretical foundation. Section 2 sets the stage and acquires the basic machinery for studying implementation. Section 3 introduces a pair of maps, called implementation maps, shows that they constitute a Galois connection, and uses this duality structure to characterize implementability.

The second part of the paper, Sections 4 to 6, illustrates the potential application of our results by developing an "abstract duality" approach to two-sided matching problems and then extending this approach to adverse-selection principal-agent problems. We must add somewhat different additional elements to our basic structure in order to obtain a matching model or a principal-agent model, but there nonetheless remains a fundamental and important connection between the two models: a profile is implementable in a principal-agent model if and only if it corresponds to a stable match in a naturally corresponding matching model. This relationship allows us to transfer arguments back and forth between the two models.

Section 4 examines two-sided matching models. We first establish the connection between implementable profiles and stable outcomes. We then use our duality results to leverage familiar existence results for matching models with a finite number of agents in order to obtain an existence result for more general models. Finally, we establish lattice results for sets of stable utility profiles.

Section 5 turns to adverse-selection principal-agent models. Our first finding is an existence result. The important step here is a strengthening of the taxation principle, showing that we can restrict attention to *implementable* tariffs when formulating the principal's maximization problem. Once again, this allows us to exploit our duality results. We next show that, unlike the quasilinear case, the solution to the principal's problem may leave slack in the participation constraint for *every* type of agent. We explore two sufficient conditions for a solution to entail a binding participation constraint. One is a *strong implementability* condition that captures the essential implication of quasilinearity in a more general form, and the other is a private values condition on the principal's payoff. In both cases, the argument exploits the lattice structure of the set of implementable utility profiles.

Section 6 considers the special case in which a single-crossing condition holds and type spaces are one dimensional. We first show that under single crossing, there exists a unique stable match that is positively assortative and (under common conditions) deterministic. With our duality results in place, the proof is a straightforward generalization of the one which yields the existence of a unique, deterministic solution to the optimal transport problem under supermodularity conditions. It then follows almost immediately from the parallels between matching and principal-agent models that an assignment is implementable if and only if it is increasing, just as in the quasilinear case.

We close the paper with a discussion of extensions. We provide further connections to the literature as we proceed.

2 Implementation

2.1 Basic Ingredients

The basic ingredients of our model are two sets, X and Y, and a function $\phi : X \times Y \times \mathbb{R} \to \mathbb{R}$ that we refer to as the *generating function*.[1] We offer two interpretations of these ingredients

Matching model. X and Y describe the possible types of two disjoint sets of agents that we refer to as buyers (X) and sellers (Y). The generating function ϕ specifies the utility frontier describing the feasible utilities 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

[1] This term is motivated by the fact that ϕ plays the same role in our analysis as the generating function of a duality plays in Penot (2000, 2010).

when matched with seller type y and providing utility v to the seller. We complete the specification of a two-sided matching model in Section 4 by specifying measures on X and Y, describing the distribution of types, and reservation utilities for the buyer and seller types.

Principal-agent model. 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. We complete the specification of an adverse-selection principal-agent model in Section 5 by specifying a utility function for the principal, her beliefs over the agent's types, and reservation utilities for the agent's types.

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

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 $\phi(x, y, v) = f(x, y) - v$. In the principal-agent model this simply means that the agent's utility function is quasilinear in the monetary transfer v. In the matching model the quasilinear case arises if both buyers and sellers have utility functions over partners and monetary transfers that are quasilinear in the transfer. While this quasilinear case is an important benchmark for our analysis, our main interest is in generating functions that are not quasilinear. In the matching context, Legros and Newman (2007, Section 5), Nöldeke and Samuelson (2015, Section 2), and Galichon, Kominers, and Weber (2016, Section 3) present a number of economic examples giving rise to non-quasilinear generating functions.[2]

In the context of the matching model, the assumption that ϕ is strictly decreasing excludes the case of *nontransferable utility* introduced in Gale and Shapley (1962), in which there is no possibility for compensatory transfers between a pair of matched agents. If the generating function is quasilinear, we have *perfectly transferable utility* as considered in Shapley and Shubik (1972), with Assumption 1 also allowing for *imperfectly transferable utility*.[3] In the context of the principal-agent model, strict monotonicity of ϕ in its third argument squares with the interpretation of this argument as a monetary transfer, while allowing for income effects. The importance of doing so in models of optimal nonlinear pricing has been emphasized in Wilson (1993, Chapter 7).

The essential implication of the full range condition in Assumption 1 is that (for example) for any agent type x and decisions y and $\tilde{y}$, one can find transfers under which the agent prefers decision y, as well as transfers under which the agent prefers decision $\tilde{y}$. Demange and Gale (1985, Section 3) discuss the importance of the full range condition in the context

[2] In the simplest case, discussed in Demange and Gale (1985), preferences are directly specified over partners and monetary transfers between matched partners. Letting $u_{xy}(t)$ denote the utility to x of being matched with y and making monetary transfer t and letting $v_{xy}(t)$ denote the corresponding utility to y of being matched with x and receiving monetary transfer t, the generating function is given by $\phi(x, y, v) = u_{xy}(v_{xy}^{-1}(v))$. This will typically be non-linear in v unless both $u_{xy}(t)$ and $v_{xy}(t)$ are linear in the transfer.

[3] Our terms for the case distinction between perfectly transferable, imperfectly transferable, and nontransferable follow (for example) Chade, Eeckhout, and Smith (2017) and Nöldeke and Samuelson (2015). Other authors (e.g. Legros and Newman, 2007) use the term nontransferable utility whenever utility is not perfectly transferable.

of the matching model. In the principal-agent model the condition ensures that the taxation principle is applicable without taking recourse to tariffs specifying infinite transfers (cf. footnote 9). The assumption that (for any given (x, y)) the generating function maps from the reals into the reals greatly simplifies the exposition, but is not essential: all of our analysis goes through if A and B are open intervals in $\mathbb{R}$ and the generating function $\phi : X \times Y \times A \to B$ satisfies the counterpart to Assumption 1.

2.2 The Inverse Generating Function

Assumption 1 ensures that for all $x \in X$, $y \in Y$ and $u \in \mathbb{R}$, there is a unique value $v \in \mathbb{R}$ satisfying $u = \phi(x, y, v)$, so that the *inverse generating function* $\psi : Y \times X \times \mathbb{R} \to \mathbb{R}$ specified as the solution to

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

is well-defined and satisfies the "reverse" inverse relationship

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

The inverse generating function inherits the properties of the generating function stated in Assumption 1: ψ is continuous, strictly decreasing in its third argument, and satisfies $\psi(y, x, \mathbb{R}) = \mathbb{R}$ for all $(y, x) \in Y \times X$.[4] Throughout the following, we will freely make use of the compactness of X and Y and the properties of the generating function ϕ and its inverse ψ without always explicitly referring to Assumption 1 or the argument in footnote 4.

In the context of the matching model the interpretation of ψ is analogous to the one given for ϕ: the utility $v = \psi(y, x, u)$ is the maximal utility a seller type y can obtain when matched with a buyer type x and providing utility u to the buyer.[5] In the principal-agent model ψ identifies the largest transfer an agent of type x can pay for the decision y while obtaining utility level u.[6] In either context, as indicated by (1)–(2), the inverse generating function contains the same information about preferences as the generating function.

2.3 Profiles, Assignments, and Implementability

Let $\mathbf{B}(X)$ denote the set of bounded functions from X to $\mathbb{R}$ and let $\mathbf{B}(Y)$ denote the set of bounded functions from Y to $\mathbb{R}$. We refer to $\boldsymbol{u} \in \mathbf{B}(X)$ and $\boldsymbol{v} \in \mathbf{B}(Y)$ as *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. We order $\mathbf{B}(X)$ and $\mathbf{B}(Y)$ with the

[4]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 $X \times Y \times \mathbb{R}$, we have $\psi(y, x, \mathbb{R}) = \mathbb{R}$ 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 real numbers is continuous if and only if its epigraph and hypograph are closed (Ferrera, 2014, Proposition 1.14, p. 5), continuity of ϕ is equivalent to continuity of ψ.

[5]Observe that in the definition of ψ the order of the first two arguments has been exchanged, so that in the 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 the quasilinear case we have $\psi(y, x, u) = g(y, x) - u$, where $g(y, x) = f(x, y)$ holds for all $(x, y) \in X \times Y$.

[6]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.

pointwise partial order inherited from the standard order $\geq$ on $\mathbb{R}$. For simplicity, we also denote these pointwise partial orders on $\mathbf{B}(X)$ and $\mathbf{B}(Y)$ by $\geq$. The join $\boldsymbol{u} \vee \boldsymbol{u}'$ and meet $\boldsymbol{u} \wedge \boldsymbol{u}'$ are respectively the pointwise maximum and minimum of the profiles $\boldsymbol{u}$ and $\boldsymbol{u}'$. With these operations the set $\mathbf{B}(X)$ (and similarly $\mathbf{B}(Y)$) is a conditionally complete lattice.[7]

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*.

We say that $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbf{B}(X) \times Y^X$ is *implementable* if there exists a profile $\boldsymbol{v} \in \mathbf{B}(Y)$ that *implements* $(\boldsymbol{u}, \boldsymbol{y})$, meaning that the conditions

$$\boldsymbol{y}(x) \in \boldsymbol{Y_v}(x) := \operatorname*{argmax}_{y \in Y} \phi(x, y, \boldsymbol{v}(y)) \tag{3}$$

$$\boldsymbol{u}(x) = \max_{y \in Y} \phi(x, y, \boldsymbol{v}(y)) \tag{4}$$

hold for all $x \in X$ (which, obviously, implies that the argmax correspondence $\boldsymbol{Y_v} : X \rightrightarrows Y$ defined in (3) is non-empty valued). Similarly, $(\boldsymbol{v}, \boldsymbol{x}) \in \boldsymbol{B}(Y) \times X^Y$ is implementable if there exists a profile $\boldsymbol{u} \in \boldsymbol{B}(X)$ implementing $(\boldsymbol{v}, \boldsymbol{x})$, meaning that

$$\boldsymbol{x}(y) \in \boldsymbol{X_u}(y) := \operatorname*{argmax}_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \tag{5}$$

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

hold 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 $(\boldsymbol{u}, \boldsymbol{y})$. We use the analogous terminology for a profile $\boldsymbol{u}$ implementing the profile $\boldsymbol{v}$ and assignment $\boldsymbol{x}$. Profiles and assignments are said to be *implementable* if there exists a profile implementing them. We let $\boldsymbol{I}(X)$ and $\boldsymbol{I}(Y)$ denote the sets of implementable profiles, so that (for example) $\boldsymbol{I}(X) = \{\boldsymbol{u} \in \boldsymbol{B}(X) | \exists \boldsymbol{v} \in \boldsymbol{B}(Y) \text{ s. t. (4) holds}\}$.

In the matching model $\boldsymbol{u}$ is a utility profile for buyers, whereas $\boldsymbol{v}$ is a utility profile for sellers. An assignment $\boldsymbol{y}$ specifies for each buyer type x a seller type $y = \boldsymbol{y}(x)$ with whom x matches; the interpretation of an assignment $\boldsymbol{x}$ is analogous.[8] In the implementation conditions (3)–(4) the utility profile $\boldsymbol{v}$ serves a dual role as a non-linear tariff: $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$ if 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 $\boldsymbol{v}$. The interpretation of conditions (5)–(6) is analogous.

In the principal-agent model $\boldsymbol{u}$ specifies a utility level for each agent type, whereas an assignment $\boldsymbol{y}$ specifies a decision for each agent type. The profile $\boldsymbol{v}$ is a non-linear tariff

[7] A lattice is conditionally complete if every non-empty subset that is bounded has both an infimum and a supremum. 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 as these two notions are equivalent in our setting: A set $\mathcal{U} \subset \boldsymbol{B}(X)$ is order bounded (i.e., there exists $\underline{\boldsymbol{u}}, \overline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ with $\underline{\boldsymbol{u}} \leq \boldsymbol{u} \leq \overline{\boldsymbol{u}}$ for all $\boldsymbol{u} \in \mathcal{U}$) only if it is norm bounded in $\boldsymbol{B}(X)$ (because $\|\boldsymbol{u}\| \leq \max\{\sup_X |\underline{\boldsymbol{u}}(x)|, \sup_X |\overline{\boldsymbol{u}}(x)|\}$ for all $\boldsymbol{u} \in \mathcal{U}$), and the converse clearly also holds, as is the case for $\mathcal{V} \subset \boldsymbol{B}(Y)$.

[8] 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 sometimes referred to as a pre-matching (Adachi, 2000) or a semi-matching (Lawler, 2001).

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. Such a tariff implements the pair $(\boldsymbol{u}, \boldsymbol{y})$ if all agent types find it optimal to choose the decisions specified in $\boldsymbol{y}$ when faced with the tariff $\boldsymbol{v}$, and $\boldsymbol{u}$ is the resulting rent function. We may think of a type assignment $\boldsymbol{x}$ as specifying for each decision y an agent type $\boldsymbol{x}(y)$ to whom the principal wants to sell decision y, as in Nöldeke and Samuelson (2007). Though the interpretation of a rent function $\boldsymbol{u}$ implementing a pair $(\boldsymbol{v}, \boldsymbol{x})$ is less obvious in the principal-agent model than in the matching model, Section 5 shows that the notion of an implementable tariff can be very helpful in the former.

Remark 1 (Implementability and Direct Mechanisms)**.** In defining implementability we have taken a nonlinear pricing (rather than a direct mechanism) approach and, in addition, have required the profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ to be both bounded. The taxation principle (e.g., Guesnerie and Laffont, 1984; Rochet, 1985) is applicable in our setting and ensures that there is no loss of generality in using a nonlinear pricing approach when studying principal-agent models. What is less obvious is that the restriction to bounded profiles is innocent, but this follows from Assumption 1.[9] Appendix A.1 provides the details. ■

2.4 Strongly Implementable Assignments

We say that an assignment is strongly implementable if it can be implemented while pegging the utility level of an arbitrary agent at an arbitrary level. Formally, an assignment $\boldsymbol{y} \in Y^X$ is *strongly implementable* if for all $(x_0, u_0) \in X \times \mathbb{R}$ there exists $\boldsymbol{u} \in \boldsymbol{B}(X)$ satisfying $\boldsymbol{u}(x_0) = u_0$ such that $(\boldsymbol{u}, \boldsymbol{y})$ is implementable. Similarly, an assignment $\boldsymbol{x} \in X^Y$ is strongly implementable if for all $(y_0, v_0) \in Y \times \mathbb{R}$ there exists $\boldsymbol{v} \in \boldsymbol{B}(Y)$ satisfying $\boldsymbol{v}(y_0) = v_0$ such that $(\boldsymbol{v}, \boldsymbol{x})$ is implementable. We will say that a profile $\boldsymbol{u}$ with $\boldsymbol{u}(x_0) = u_0$ satisfies the initial condition (x_0, u_0) and that a profile $\boldsymbol{v}$ with $\boldsymbol{v}(y_0) = v_0$ satisfies the initial condition (y_0, v_0).

With a quasilinear generating function every implementable assignment is strongly implementable, so that the distinction between these two concepts is moot. This follows from the translational invariance of the incentive constraints under quasilinearity: $\boldsymbol{u}(x) = f(x, \boldsymbol{y}(x)) - \boldsymbol{v}(\boldsymbol{y}(x)) = \max_{y \in Y} [f(x, y) - \boldsymbol{v}(y)]$ implies $\boldsymbol{u}(x) - t = f(x, \boldsymbol{y}(x)) - (\boldsymbol{v}(\boldsymbol{y}(x)) + t) = \max_{y \in Y} [f(x, y) - (\boldsymbol{v}(y) + t)]$ for all $x \in X$ and $t \in \mathbb{R}$, so that by choosing the constant t appropriately a tariff $\boldsymbol{v}$ implementing an assignment $\boldsymbol{y}$ can be adjusted to satisfy any given initial condition while continuing to implement $\boldsymbol{y}$ (with an analogous argument applying to implementable $\boldsymbol{x} \in X^Y$).

In general, the implementability of an assignment does not imply its strong implementability. It will become clear as we progress that this causes some salient differences between the quasilinear and the general case. For example, if every implementable profile is strongly implementable, then—just as in the quasilinear case—the participation constraint must be binding for some type of agent in a solution to the principal-agent model considered in Section 5, whereas this property may fail otherwise (cf. Proposition 10 and Example 2 in Section 5.3). Sections 6.2 and 7.1 identify circumstances in which all implementable

[9] In the absence of the full range condition from Assumption 1 this conclusion may fail. To see this, it suffices to consider an incentive compatible allocation in which type x obtains utility u from choosing y, but there exists y' such that $\lim_{v \to \infty} \phi(x, y', v) \geq u$. Then, no matter what transfer $\boldsymbol{v}(y') \in \mathbb{R}$ is specified, type x will prefer to choose y' rather than y.

profiles are strongly implementable, ensuring that an important structural property of the quasilinear case is preserved, even though the generating function is not quasilinear.

3 Duality

In this section we characterize implementable profiles and assignments. Section 3.1 introduces a pair of functions between sets of profiles that we refer to as implementation maps. Section 3.1 shows that the implementation maps are a Galois connection between the sets of profiles $\mathbf{B}(X)$ and $\mathbf{B}(Y)$. Equivalently, these maps are dualities in the sense of Penot (2010) that are dual to each other, motivating the titles of this section and the paper. Section 3.2 uses the structure of the implementation maps to characterize implementable profiles. Building on these results, Section 3.3 characterizes implementable assignments and Section 3.4 establishes some key properties of sets of implementable profiles.

3.1 Implementation Maps

Consider any profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$. As X and Y are compact and ϕ is continuous, setting $\boldsymbol{u}(x) = \sup_{y \in Y} \phi(x, y, \boldsymbol{v}(y))$ for all $x \in X$ results in a bounded profile $\boldsymbol{u} \in \boldsymbol{B}(X)$. Together with a similar argument for $\boldsymbol{v}(y) = \sup_{x \in X} \psi(y, x, \boldsymbol{u}(x))$, this ensures that the implementation maps $\Phi : \mathbf{B}(Y) \to \mathbf{B}(X)$ and $\Psi : \mathbf{B}(X) \to \mathbf{B}(Y)$ obtained by setting

$$\Phi\boldsymbol{v}(x) = \sup_{y \in Y} \phi(x, y, \boldsymbol{v}(y)) \quad \forall x \in X \tag{7}$$

$$\Psi\boldsymbol{u}(y) = \sup_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \quad \forall y \in Y \tag{8}$$

are well-defined. Appendix A.2 proves:

Lemma 1. *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 continuous and map bounded sets into bounded sets.*

Many useful properties of the implementation maps can be obtained as a simple consequence of the fact that Φ and Ψ are a Galois connection (Birkhoff, 1995, p. 124) between the sets $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$. That is,

$$\boldsymbol{u} \;\geq\; \Phi\boldsymbol{v} \iff \boldsymbol{v} \;\geq\; \Psi\boldsymbol{u} \tag{9}$$

holds for all $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$.[10]

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

[10] There is an alternative definition of a Galois connection in which the second inequality in (9) is reversed (Davey and Priestley, 2002, Chapter 7). This difference is not essential, as equivalence between the two definitions can be restored by replacing $\boldsymbol{B}(Y)$ with its order dual. The kind of Galois connection we consider is sometimes referred to as an antitone Galois connection.

Proof. To obtain (9) 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 (7), 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, the fourth is by the definition of the supremum, and the fifth holds by the definition of $\Psi\boldsymbol{u}$ in (8). ■

To interpret Proposition 1, consider the matching context. 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 explanation motivates the other direction of the equivalence in (9).

The first three statements in the following corollary are standard implications of the fact that Φ and Ψ are a Galois connection. Our terms for these follow Davey and Priestley (2002, p. 159).[11] The fourth statement in the corollary observes that the implementation maps Φ and Ψ are dualities that are dual to each other. We use the term duality as does Penot (2010, Definition 1, page 505), who defines a duality as a map between two partially ordered sets with the property that for any subset of the domain which has an infimum, the image of the infimum of that set is the supremum of its image. We say that the implementation maps are dual to each other if

$$\Phi\boldsymbol{v} = \inf\{\boldsymbol{u} | \boldsymbol{v} \geq \Psi\boldsymbol{u}\} \text{ and } \Psi\boldsymbol{u} = \inf\{\boldsymbol{v} | \boldsymbol{u} \geq \Phi\boldsymbol{v}\}$$

holds for all $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$.[12] As the proof of the corollary is straightforward and makes clear how one may infer properties of the implementation maps from the mere knowledge that they are a Galois connection, we include it here.

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

[1.1] satisfy the cancellation rule, that is, for all $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$:

$$\boldsymbol{v} \geq \Psi\Phi\boldsymbol{v} \text{ and } \boldsymbol{u} \geq \Phi\Psi\boldsymbol{u}; \tag{10}$$

[11] As noted in Birkhoff (1995, Section 5.8), the properties stated in Corollary 1.1–1.2 below are in fact equivalent to (9) and are sometimes taken to be the definition of a Galois connection (e.g., Singer, 1997, Definition 5.3 and Remark 5.6). See also the original definition of a Galois connection in Ore (1944).

[12] In his monograph on abstract convex analysis, Singer (1997, Definition 5.1, p. 172) defines a duality as a map between *complete lattices* with the property that the image of the infimum of any set is the supremum of the image of that set. Penot's definition provides the obvious generalization to the situation under consideration here in which $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$ are lattices, but are not complete. Our usage of the term dual map is similarly adapted from Singer (1997, Definition 5.2, p. 176).

[1.2] are order reversing, that is, for all $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \boldsymbol{B}(X)$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \boldsymbol{B}(Y)$:

$$\boldsymbol{v}_1 \geq \boldsymbol{v}_2 \Rightarrow \Phi\boldsymbol{v}_1 \leq \Phi\boldsymbol{v}_2 \text{ and } \boldsymbol{u}_1 \geq \boldsymbol{u}_2 \Rightarrow \Psi\boldsymbol{u}_1 \leq \Psi\boldsymbol{u}_2; \tag{11}$$

[1.3] satisfy the semi-inverse rule, that is, for all $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$:

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

[1.4] and are dualities that are dual to each other.

Proof. We use (9) to establish (10)–(12). In each case we prove one of the two statements; the other statement follows by an analogous argument. First, for any $\boldsymbol{v} \in \boldsymbol{B}(Y)$ we trivially have $\Phi\boldsymbol{v} \geq \Phi\boldsymbol{v}$, so that setting $\boldsymbol{u} = \Phi\boldsymbol{v}$ in (9) yields (10). Second, let $\boldsymbol{v}_1 \geq \boldsymbol{v}_2$. By (10) we have $\boldsymbol{v}_2 \geq \Psi\Phi\boldsymbol{v}_2$ and thus $\boldsymbol{v}_1 \geq \Psi\Phi\boldsymbol{v}_2$. Applying (9) with $\boldsymbol{v} = \boldsymbol{v}_1$ and $\boldsymbol{u} = \Phi\boldsymbol{v}_2$ then gives the consequent of (11). Third, (10) gives $\boldsymbol{v} \geq \Psi\Phi\boldsymbol{v}$. Applying (11) 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 the reverse inequality and hence (12), notice that for every $\boldsymbol{v} \in \boldsymbol{B}(Y)$ 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 (9) yields the reverse inequality $\Phi\boldsymbol{v} \geq \Phi\Psi\Phi\boldsymbol{v}$.

To confirm that Φ is a duality (with Ψ analogous), let $\underline{\boldsymbol{v}}$ be the infimum of some set $\mathcal{V} \subset \boldsymbol{B}(Y)$. Corollary 1.2 implies that $\Phi\underline{\boldsymbol{v}}$ then is an upper bound of $\Phi\mathcal{V}$. Let $\overline{\boldsymbol{u}}$ be any upper bound of $\Phi\mathcal{V}$. By (9) we then have $\boldsymbol{v} \geq \Psi\overline{\boldsymbol{u}}$ for all $\boldsymbol{v} \in \mathcal{V}$, implying $\underline{\boldsymbol{v}} \geq \Psi\overline{\boldsymbol{u}}$. Applying (9) again, this yields $\overline{\boldsymbol{u}} \geq \Phi\underline{\boldsymbol{v}}$, showing that $\Phi\underline{\boldsymbol{v}}$ is the supremum of $\Phi\mathcal{V}$. Finally, (9) implies $\{\boldsymbol{u}|\boldsymbol{v} \geq \Psi\boldsymbol{u}\} = \{\boldsymbol{u}|\boldsymbol{u} \geq \Phi\boldsymbol{v}\}$, so that $\inf\{\boldsymbol{u}|\boldsymbol{v} \geq \Psi\boldsymbol{u}\} = \inf\{\boldsymbol{u}|\boldsymbol{u} \geq \Phi\boldsymbol{v}\} = \Phi\boldsymbol{v}$. An analogous argument establishes $\Psi\boldsymbol{u} = \inf\{\boldsymbol{v}|\boldsymbol{u} \geq \Phi\boldsymbol{v}\}$, so that Φ and Ψ are dual to each other. ■

To provide some interpretation for (10)–(12) we focus on the first statement in each case and consider the principal-agent model. The order-reversal property (Corollary 1.2) simply asserts that all agent types are better off when the prices specified by the tariff are low rather than high. Intuitively, the tariff $\Psi\Phi\boldsymbol{v}$ appearing in the cancellation rule (Corollary 1.1) specifies for each decision $y \in Y$ the highest possible payment compatible with providing some agent type with the same utility from choosing y than the one which the agent type obtains from maximizing against the tariff $\boldsymbol{v}$, thereby making it an "envelope tariff" (see footnote 13 below). The assertion of the cancellation rule then is that the envelope tariff $\Psi\Phi\boldsymbol{v}$ obtained from the tariff $\boldsymbol{v}$ specifies payments no higher than the original tariff $\boldsymbol{v}$. Finally, the semi-inverse rule (Corollary 1.3) is the assertion that the inequality from the cancellation rule turns into an equality when the original tariff $\boldsymbol{v}$ solves the problem of specifying the highest possible payments compatible with providing some agent type with the utility level $\boldsymbol{u}(x)$ for each decision y.

Remark 2 (Quasilinearity and Generalized Conjugate Duality)**.** In the quasilinear case the definitions of the implementation maps in (7) and (8) reduce to

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

where $g(y, x) = f(x, y)$ holds for all $(x, y) \in X \times Y$ (cf. footnote 5). In this case $\Phi\boldsymbol{v}$ is a familiar object, namely the f-conjugate of $\boldsymbol{v}$, and $\Psi\boldsymbol{u}$ is the g-conjugate of $\boldsymbol{u}$ (cf. Ekeland, 2010, Section 3.2). The first three properties noted in Corollary 1 may then be viewed as generalizing corresponding properties from the theory of (generalized) 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, 2010, Section 3.4).[13] Martinez-Legaz and Singer (1990, 1995) offer other illustrations of how results for abstract dualities specialize to familiar results from conjugate duality when the generating function is quasilinear. ■

Remark 3 (Dualities and Galois Connections)**.** It is noteworthy that the statement in Corollary 1.4 is not only implied by Proposition 1 but implies it. Therefore, all the consequences of Proposition 1 that we employ in our subsequent analysis, including Corollary 1.1–1.3, may be viewed as implications of the duality structure of the implementation maps. The proof is straightforward (cf. Singer, 1997, p. 179): Let $\boldsymbol{u} \geq \Phi\boldsymbol{v}$. Then $\Psi\boldsymbol{u} \leq \Psi\Phi\boldsymbol{v} \leq \inf\{\tilde{\boldsymbol{v}} | \Phi\tilde{\boldsymbol{v}} \leq \Phi\boldsymbol{v}\} \leq \boldsymbol{v}$, where the first inequality follows from the order-reversing property of the duality Ψ, the equality follows from the fact that Ψ and Φ are dual, and the final inequality from the definition of the infimum. This gives one of the implications of (9); the other is analogous. ■

3.2 Implementable Profiles

Comparing the implementation condition (4) and the definition of the implementation map Φ in (7) it is clear that $\boldsymbol{v} \in \boldsymbol{B}(Y)$ implements $\boldsymbol{u} \in \boldsymbol{B}(X)$ if and only if $\boldsymbol{u} = \Phi\boldsymbol{v}$ holds and, in addition, the suprema in (7) are attained for all $x \in X$, that is, the argmax correspondence $\boldsymbol{Y_v}$ defined in (3) is non-empty valued. Consequently, the set of implementable profiles $\boldsymbol{I}(X)$ is contained in the image of the implementation map Φ. Similarly, $\boldsymbol{I}(Y) \subseteq \Psi\boldsymbol{B}(X)$ holds.

The following proposition shows that the reverse set inclusion also holds. Hence, the images of the implementation maps are precisely the sets of implementable profiles. In the course of proving this result, it is straightforward to also show that every implementable profile is continuous.[14] Let $\boldsymbol{C}(X) \subseteq \boldsymbol{B}(X)$ denote the set of continuous (and hence necessarily bounded, since X is compact) functions from X to $\mathbb{R}$, with $\boldsymbol{C}(Y)$ analogous. Appendix A.3 shows:

Proposition 2. *Let Assumption 1 hold. A profile is implementable if and only if it is in the image of the relevant implementation map. Further, every implementable profile is continuous. That is,*

$$\boldsymbol{I}(X) = \Phi\boldsymbol{B}(Y) \subseteq \boldsymbol{C}(X) \quad \textit{and} \quad \boldsymbol{I}(Y) = \Psi\boldsymbol{B}(X) \subseteq \boldsymbol{C}(Y). \tag{13}$$

[13] In convex analysis, the counterpart of $\Psi\Phi\boldsymbol{v}$ is also referred to as the convex envelope of $\boldsymbol{v}$, and is the greatest convex minorant of $\boldsymbol{v}$ (Galichon, 2016, Proposition D.12, p. 157). An analogous property holds here. First, from the cancellation property, $\Psi\Phi\boldsymbol{v}$ is a minorant of $\boldsymbol{v}$. Second, consider $\boldsymbol{u}$ satisfying $\Psi\boldsymbol{u} \leq \boldsymbol{v}$. Applying the order reversal property twice yields $\Psi\Phi\Psi\boldsymbol{u} \leq \Psi\Phi\boldsymbol{v}$ and therefore, from the semi-inverse rule $\Psi\boldsymbol{u} \leq \Psi\Phi\boldsymbol{v}$.

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

The first step in the proof of Proposition 2 (Lemma 8 in Appendix A.3) shows that every lower semicontinuous profile implements its image under the relevant implementation map and that this image is continuous. The proof of the proposition is then completed by showing that for any profile, its image under the relevant implementation map is the same as the image of its lower semicontinuous hull.

The continuity (and hence lower semicontinuity) of implementable profiles ensures that every implementable profile implements its image under the relevant implementation map and that the attendant argmax correspondences are nicely behaved. For later reference we state this observation, which is a direct consequence of Weierstrass' extreme value theorem and Berge's maximum theorem (cf. the proof of Lemma 8 in Appendix A.3), as a corollary.

Corollary 2. *Let Assumption 1 hold and let* $\boldsymbol{v} \in \boldsymbol{I}(Y)$ *and* $\boldsymbol{u} \in \boldsymbol{I}(X)$. *Then* $\boldsymbol{v}$ *implements* $\Phi\boldsymbol{v}$ *and the argmax correspondence* $\boldsymbol{Y}_{\boldsymbol{v}}$ *is nonempty- and compact-valued and upper hemicontinuous. Analogously,* $\boldsymbol{u}$ *implements* $\Psi\boldsymbol{u}$ *and the argmax correspondence* $\boldsymbol{X}_{\boldsymbol{u}}$ *is nonempty- and compact-valued and upper hemicontinuous.*

Combining Proposition 2 with the semi-inverse rule from Corollary 1.3 yields a characterization of implementable profiles:

Proposition 3. *Let Assumption 1 hold.*
[3.1] $\boldsymbol{u} \in \boldsymbol{B}(X)$ *is implementable if and only if* $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$.
[3.2] $\boldsymbol{v} \in \boldsymbol{B}(Y)$ *is implementable if and only if* $\boldsymbol{v} = \Psi\Phi\boldsymbol{v}$.

Proof. We prove Proposition 3.1; 3.2 is analogous. First, we show

$$\boldsymbol{u} \in \Phi\boldsymbol{B}(Y) \iff \boldsymbol{u} = \Phi\Psi\boldsymbol{u}.$$

To establish this, recall that the images of the implementation maps are contained in the sets $\boldsymbol{B}(Y)$, resp. $\boldsymbol{B}(X)$. Hence, $\Psi\boldsymbol{u} \in \boldsymbol{B}(Y)$ so that $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$ implies $\boldsymbol{u} \in \Phi\boldsymbol{B}(Y)$. Conversely, if $\boldsymbol{u} \in \Phi\boldsymbol{B}(Y)$, then there exists $\boldsymbol{v} \in \boldsymbol{B}(Y)$ such that $\boldsymbol{u} = \Phi\boldsymbol{v}$, and hence (by Corollary 1.3) we have $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$. Second, using the above equivalence, Proposition 3.1 follows from the equality $\boldsymbol{I}(X) = \Phi\boldsymbol{B}(Y)$ in Proposition 2. ■

The first part of the above proof of Proposition 3 uses nothing but the semi-inverse rule from Corollary 1.3 to establish that the fixed point condition $u = \Phi\Psi u$ characterizes the image of the map Φ and, analogously, the fixed point condition $v = \Psi\Phi v$ characterizes the image of the map Ψ. Hence, this holds for any Galois connection (see also Singer, 1997, Corollary 5.6). An appeal to Proposition 2 completes the proof of Proposition 3 by providing the connection between images of the implementation maps and the sets of implementable profiles. This second step uses the special structure of our duality provided by Assumption 1. Assumption 1 thus 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 a characterization of implementable profiles.

It is a straightforward implication of Propositions 2 and 3 that the sets of implementable profiles coincide with the images of the sets of implementable profiles under the relevant implementation map. That is, we have

$$\boldsymbol{I}(X) = \Phi\boldsymbol{I}(Y) \quad \text{and} \quad \boldsymbol{I}(Y) = \Psi\boldsymbol{I}(X). \tag{14}$$

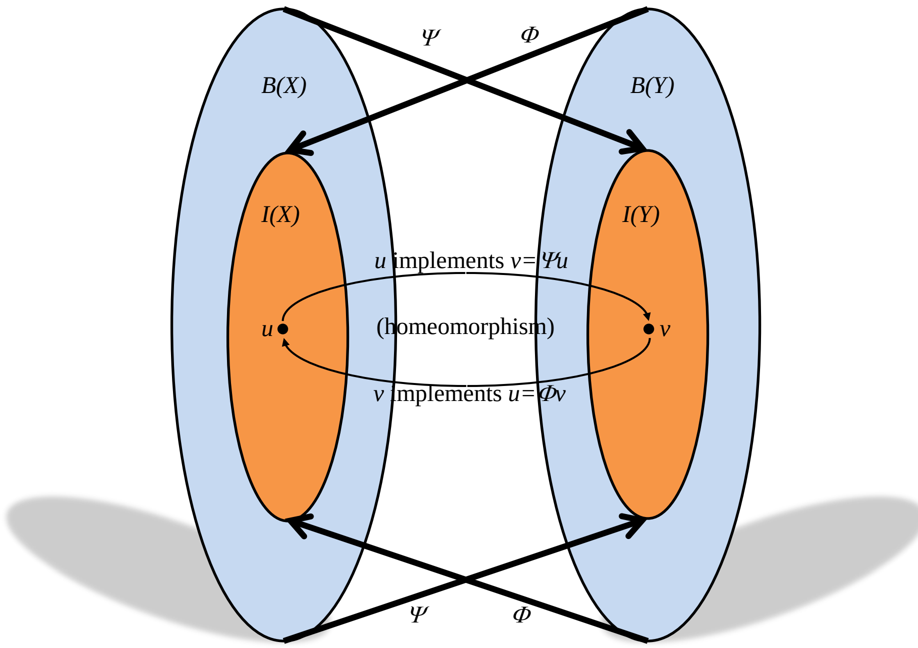

Figure 1: Illustration of the implementation maps. The implementation map Φ maps the set of bounded profiles $\boldsymbol{B}(Y)$ into the set of implementable profiles $\boldsymbol{I}(X)$ (and Ψ maps the set of bounded profiles $B(X)$ into the set of implementable profiles $\boldsymbol{I}(Y)$). The maps Φ and Ψ are continuous inverse bijections on the sets of implementable profiles $\boldsymbol{I}(X)$ and $\boldsymbol{I}(Y)$ with profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ in these sets satisfying $\boldsymbol{u} = \Phi\boldsymbol{v} \iff \boldsymbol{v} = \Psi\boldsymbol{u}$ and implementing each other.

To see this, consider the case of an implementable profile $\boldsymbol{u} \in \boldsymbol{I}(X)$. By Proposition 3.1 the profile $\boldsymbol{u}$ is the image of the profile $\boldsymbol{v} = \Psi\boldsymbol{u}$ under the implementation map Φ, with the profile $\boldsymbol{v} = \Psi\boldsymbol{u}$ being implementable by Proposition 2. This establishes $\boldsymbol{I}(X) \subseteq \Phi\boldsymbol{I}(Y)$. The reverse set inclusion is immediate from $\boldsymbol{I}(X) = \Phi\boldsymbol{B}(Y)$ (Proposition 2) and $\boldsymbol{I}(Y) \subset \boldsymbol{B}(Y)$.

In addition, the implementation maps Φ and Ψ are inverse bijections between $\boldsymbol{I}(X)$ and $\boldsymbol{I}(Y)$. That is, we have

$$\boldsymbol{u} = \Phi\boldsymbol{v} \iff \boldsymbol{v} = \Psi\boldsymbol{u}, \quad \text{for all } \boldsymbol{u} \in \boldsymbol{I}(X) \text{ and } \boldsymbol{v} \in \boldsymbol{I}(Y). \tag{15}$$

To see this, suppose $\boldsymbol{u} = \Phi\boldsymbol{v}$ holds for some implementable profiles $\boldsymbol{u}$ and $\boldsymbol{v}$. Applying the implementation map Ψ to both sides of this equality yields $\Psi\boldsymbol{u} = \Psi\Phi\boldsymbol{v}$. As $\boldsymbol{v}$ is implementable, we have $\Psi\Phi\boldsymbol{v} = \boldsymbol{v}$ from Proposition 3.2, implying $\Psi\boldsymbol{u} = \boldsymbol{v}$. An analogous argument works in reverse, delivering (15). As the implementation maps are continuous (Lemma 1) the implementation maps are thus homeomorphisms between the sets of implementable profiles.

Because implementable profiles implement their images (Corollary 2) the equalities in (14) can be rephrased as the statement that every implementable profile can be implemented by an implementable profile, with (15) then identifying the unique such profile. In summary, every implementable profile $\boldsymbol{u}$ is implemented by the implementable profile $\boldsymbol{v} = \Psi\boldsymbol{u}$ and, dually, every implementable profile $\boldsymbol{v}$ is implemented by the implementable profile $\boldsymbol{u} = \Phi\boldsymbol{v}$. Figure 1 illustrates these observations and puts them into the context provided by Proposition 2.

Remark 4 (Implementable Profiles in the Quasilinear Case)**.** Following up on Remark 2, we note that in the quasilinear case Proposition 3 is the statement that a profile is implementable if and only if it is its own generalized biconjugate (Ekeland, 2010, Corollary 12). ■

3.3 Implementable Assignments

Consider an implementable $(\boldsymbol{u}, \boldsymbol{y})$. As we have just argued, Proposition 3 ensures that $\boldsymbol{u}$ can then be implemented by an implementable profile, namely $\Psi\boldsymbol{u}$. The following result formalizes this observation and, in addition, addresses the question whether $\Psi\boldsymbol{u}$ also implements $\boldsymbol{y}$ (and the analogous question for implementable $(\boldsymbol{v}, \boldsymbol{x})$). We obtain an affirmative answer. Therefore, not only all implementable profiles but also all implementable assignments can be implemented by implementable profiles. The proof is in Appendix A.4.

Corollary 3. *Let Assumption 1 hold.*

*[3.1] The pair $(\boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{B}(X) \times Y^X$ is implementable if and only if both (*i*) $\boldsymbol{u}$ implements $\Psi\boldsymbol{u}$ and (*ii*) $\Psi\boldsymbol{u}$ implements $(\boldsymbol{u}, \boldsymbol{y})$.*

*[3.2] The pair $(\boldsymbol{v}, \boldsymbol{x}) \in \boldsymbol{B}(Y) \times X^Y$ is implementable if and only if both (*i*) $\boldsymbol{v}$ implements $\Phi\boldsymbol{v}$ and (*ii*) $\Phi\boldsymbol{v}$ implements $(\boldsymbol{v}, \boldsymbol{x})$.*

Corollary 3 provides a conceptually simple test for deciding whether a pair $(\boldsymbol{v}, \boldsymbol{y})$ or $(\boldsymbol{u}, \boldsymbol{x})$ featuring a continuous profile is implementable:[15] Given any such pair, determine the profile implemented by the profile in the pair. If that profile implements the pair, then the pair is implementable; otherwise it is not.[16]

Recall that $\boldsymbol{Y}_{\boldsymbol{v}}$ and $\boldsymbol{X}_{\boldsymbol{u}}$, defined in equations (3) and (5), denote the argmax correspondences from the maximization problems that define implementability. From Corollary 3, an assignment is implementable if and only if it is a selection from one of these correspondences for an implementable profile. To make progress in characterizing implementable assignments it is therefore natural to consider the structure of $\boldsymbol{Y}_{\boldsymbol{v}}$ and $\boldsymbol{X}_{\boldsymbol{u}}$ for implementable $\boldsymbol{u}$ and $\boldsymbol{v}$.

Given any pair of profiles $(\boldsymbol{u}, \boldsymbol{v})$ let

$$\begin{aligned} \Gamma_{\boldsymbol{u},\boldsymbol{v}} &= \{(x, y) \in X \times Y \mid \boldsymbol{u}(x) = \phi(x, y, \boldsymbol{v}(y))\} \\ &= \{(x, y) \in X \times Y \mid \boldsymbol{v}(y) = \psi(y, x, \boldsymbol{u}(x))\}, \end{aligned} \tag{16}$$

where the second equality holds by definition of the inverse generating function ψ. If $\boldsymbol{v}$ implements $\boldsymbol{u}$, then $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ coincides with the graph of $\boldsymbol{Y}_{\boldsymbol{v}}$, i.e., $\{(x, y) \in X \times Y | \boldsymbol{v}(y) = \max_{\tilde{x} \in X} \psi(y, \tilde{x}, \boldsymbol{u}(\tilde{x})\} = \{(x, y) \in X \times Y \mid \boldsymbol{v}(y) = \psi(y, x, \boldsymbol{u}(x))\}$. Similarly, if $\boldsymbol{u}$ implements $\boldsymbol{v}$, the equality in the second line indicates that $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ coincides with the graph of $\boldsymbol{X}_{\boldsymbol{u}}$. For the special case in which the profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other, the graphs of both $\boldsymbol{X}_{\boldsymbol{u}}$ and $\boldsymbol{Y}_{\boldsymbol{v}}$ thus coincide with $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$, indicating that the two argmax correspondences are inverse to each other. This proves:

[15] If the profile in the pair is not continuous, then it is of course immediate from Proposition 2 that the pair is not implementable.

[16] In the quasilinear case, Corollary 3.1 is equivalent to the statement that a profile-assignment pair $(\boldsymbol{u}, \boldsymbol{y})$ is implementable if and only if it is implemented by the generalized conjugate of $\boldsymbol{u}$. As discussed in Basov (2006, p. 136 and p. 142) the latter result is the essence of the implementability criterion for the quasilinear case provided by Carlier (2002, Proposition 1).

Lemma 2. *Let Assumption 1 hold and suppose that $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other. The argmax correspondences $\boldsymbol{X}_{\boldsymbol{u}}$ and $\boldsymbol{Y}_{\boldsymbol{v}}$ are inverses and their graphs coincide with $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$, i.e., they satisfy*

$$\hat{x} \in \boldsymbol{X}_{\boldsymbol{u}}(\hat{y}) \iff \hat{y} \in \boldsymbol{Y}_{\boldsymbol{v}}(\hat{x}) \iff (\hat{x}, \hat{y}) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}. \tag{17}$$

Lemma 2 indicates that the inverse relationship (15) between profiles that implement each other extends to the argmax correspondences associated with these two profiles.[17] Making use of Corollary 3 this observation leads to the following characterization of implementable assignments.

Proposition 4. *Let Assumption 1 hold.*

[4.1] An assignment $\boldsymbol{y} \in Y^X$ is implementable if and only if there exist profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ that implement each other with $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ containing the graph of $\boldsymbol{y}$, i.e.,

$$(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}} \quad \text{for all } x \in X.$$

[4.2] An assignment $\boldsymbol{x} \in X^Y$ is implementable if and only if there exist profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ that implement each other with $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ containing the graph of $\boldsymbol{x}$, i.e.,

$$(\boldsymbol{x}(y), y) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}} \quad \text{for all } y \in Y.$$

Proof. We prove Proposition 4.1; 4.2 is analogous. First, suppose the profiles $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$ implement each other and let $\boldsymbol{y} \in Y^X$ satisfy $(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ for all $x \in X$. Then it follows from (17) in Lemma 2 that for all $x \in X$, we have $\boldsymbol{y}(x) \in \boldsymbol{Y}_{\boldsymbol{v}}(x)$. Hence $\boldsymbol{v}$ implements $\boldsymbol{y}$ (cf. (3)) and $\boldsymbol{y}$ is therefore implementable. Second, suppose that $\boldsymbol{y} \in Y^X$ is implementable, i.e., suppose there exists $\boldsymbol{u}$ such that $(\boldsymbol{u}, \boldsymbol{y})$ is implementable. Let $\boldsymbol{v} = \Psi\boldsymbol{u}$. Then, from Corollary 3.1, $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other and $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$. From (3), we then have that for all $x \in X$, $\boldsymbol{y}(x) \in \boldsymbol{Y}_{\boldsymbol{v}}$. Using Lemma 2, it then follows that for all $x \in X$, we have $(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$, finishing the proof. ■

Remark 5 (Implementable Assignments and Strong Implementability)**.** In the quasilinear case an assignment is implementable if and only if it satisfies the cyclical-monotonicity condition identified in Rochet (1987, Theorem 1) (see Vohra (2011, Chapter 4) for a discussion). Importantly, and in contrast to the characterization result in Proposition 4, cyclical monotonicity is a condition on assignments that does not involve any profiles and therefore can be verified directly.[18] In general, the existence of implementable assignments that are not strongly implementable precludes any hope of obtaining a condition that allows to verify the implementability of an assignment without considering the associated profiles. On the other hand, if it is known that (as in the quasilinear case) all implementable assignments are strongly implementable, a sharper characterization of implementable assignments might be possible. Section 6 provides an illustration. ■

[17] The counterpart of Lemma 2 in the quasilinear case is the following: if $\boldsymbol{u}$ and $\boldsymbol{v}$ are each others' conjugates, then the graphs of both of their subdifferentials coincide with the set of points for which equality holds in the Fenchel inequality (cf. Ekeland, 2010, Corollary 13).

[18] In essence, Rochet's proof of his Theorem 1 shows how to construct $\boldsymbol{u}$ and $\boldsymbol{v}$ satisfying the sufficient conditions in Proposition 4 if the assignment is cyclical monotone, and also shows that doing so is impossible if cyclical monotonicity fails.

Remark 6 (Another Characterization of Implementable Profiles)**.** Proposition 4 characterizes implementable assignments in terms of the argmax correspondences $\boldsymbol{X_u}$ and $\boldsymbol{Y_v}$. Implementable profiles can be characterized in a analogous way. In particular, suppose the profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is implementable. Then $\boldsymbol{v}$ implements and is implemented by $\boldsymbol{u} = \Phi\boldsymbol{v}$ (Corollary 3), implying that both $X_{\boldsymbol{u}}$ and $Y_{\boldsymbol{v}}$ are nonempty valued. Further, from Lemma 2 the correspondences are inverses of each other, and hence must be onto. Appendix A.5 shows that the converse holds as well. Hence, for $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$, we have

$$\boldsymbol{u} \in \boldsymbol{I}(X) \iff X_{\boldsymbol{u}} \text{ is nonempty} - \text{valued and onto} \tag{18}$$
$$\boldsymbol{v} \in \boldsymbol{I}(Y) \iff Y_{\boldsymbol{v}} \text{ is nonempty} - \text{valued and onto.} \tag{19}$$

■

3.4 Sets of Implementable Profiles

We use $\mathcal{U}_{\boldsymbol{y}}$ to denote the subset of implementable profiles $\boldsymbol{I}(X)$ for which $(\boldsymbol{u}, \boldsymbol{y})$ is implementable and define $\mathcal{V}_{\boldsymbol{x}}$ analogously:

$$\mathcal{U}_{\boldsymbol{y}} = \{\boldsymbol{u} \in \boldsymbol{I}(X) : (\boldsymbol{u}, \boldsymbol{y}) \text{ is implementable}\} \tag{20}$$
$$\mathcal{V}_{\boldsymbol{x}} = \{\boldsymbol{v} \in \boldsymbol{I}(Y) : (\boldsymbol{v}, \boldsymbol{x}) \text{ is implementable}\}. \tag{21}$$

We will sometimes refer to these sets as the set of profiles compatible with $\boldsymbol{y}$, resp. with $\boldsymbol{x}$.

3.4.1 Metric Structure

The following corollary establishes properties of sets of implementable profiles that play a key role throughout our study of matching and principal-agent models in the following sections.

Corollary 4. *Let Assumption 1 hold. Then,*
[4.1] $\boldsymbol{I}(X)$ *is closed and so is* $\mathcal{U}_{\boldsymbol{y}}$ *for all* $\boldsymbol{y} \in Y^X$.
[4.2] If $\mathcal{U} \subset \boldsymbol{I}(X)$ *is bounded, then it is equicontinuous.*
[4.3] If $\mathcal{U} \subset \boldsymbol{I}(X)$ *is closed and bounded, then it is compact.*
Analogously, $\boldsymbol{I}(Y)$ *and* $\mathcal{V}_{\boldsymbol{x}}$ *are closed, if* $\mathcal{V} \subset \boldsymbol{I}(Y)$ *is bounded, then it is equicontinuous, and if it is closed and bounded, then it is compact.*

Appendix A.6 contains the proof. First, we use the duality results of Corollary 3 to show that for any converging sequence of profiles in $\boldsymbol{I}(X)$, there exists a converging sequence of profiles in $\boldsymbol{I}(Y)$ that implement the former sequence. A limiting implementability relationship then follows from the continuity of the implementation map Φ (Lemma 1), allowing us to conclude that $\boldsymbol{I}(X)$ is closed. The result about the closedness of $\mathcal{U}_{\boldsymbol{y}}$ follows by the same argument. Next, Corollary 4.2 is established by using Lemma 1 and Corollary 3 to conclude that any bounded set $\mathcal{U} \subset I(X)$ is implemented by a bounded set $\mathcal{V}$ of profiles (namely the image of the set $\mathcal{U}$ under the implementation map Ψ). This combines with the incentive constraints for implementation and the continuity of ϕ to imply the equicontinuity result. Finally, Corollary 4.3 follows from Corollary 4.2 by applying the Arzela-Ascoli theorem.

3.4.2 Order Structure

The fact that the implementation maps are dualities implies that the sets of implementable profiles are join semi-sublattices of the lattices of profiles: If, say, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are implementable, then (Corollary 3.1) there exist $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in $\boldsymbol{I}(Y)$ implementing them, i.e., such that $\boldsymbol{u}_1 = \Phi\boldsymbol{v}_1$ and $\boldsymbol{u}_2 = \Phi\boldsymbol{v}_2$. Because Φ is a duality (Corollary 1.4), we immediately have $\Phi(\boldsymbol{v}_1 \wedge \boldsymbol{v}_2) = \boldsymbol{u}_1 \vee \boldsymbol{u}_2$. The implementability of $\boldsymbol{u}_1 \vee \boldsymbol{u}_2$ then follows from Proposition 2.

If it were the case that for profiles $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{v}_1$, and $\boldsymbol{v}_2$ with the properties from the previous paragraph we also have $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2) = \boldsymbol{u}_1 \wedge \boldsymbol{u}_2$, then we could conclude that $\boldsymbol{I}(X)$ is a sublattice of $\boldsymbol{B}(X)$. Indeed, this condition is not only sufficient but also necessary for $\boldsymbol{I}(X)$ to be a sublattice: An argument analogous to the one in the preceding paragraph tells us that $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ implements $\boldsymbol{v}_1 \vee \boldsymbol{v}_2$. Therefore, Corollary 3.2 implies that $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ is implementable if and only if it is implemented by $\boldsymbol{v}_1 \vee \boldsymbol{v}_2$.

Alas, $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2) = \boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ need not hold even if the generating function is quasilinear:

Example 1. Let $X = \{1, 2, 3\}$ and $Y = \{1, 2\}$ and let the generating function be given by

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

for $x \in X$. The profiles $\boldsymbol{u}_1 = (1, 1, 1)$ and $\boldsymbol{v}_1 = (0, 4)$ as well as the profiles $\boldsymbol{u}_2 = (0, 1, 2)$ and $\boldsymbol{v}_2 = (1, 3)$ implement each other. Further, $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2 = (0, 1, 1)$, $\boldsymbol{v}_1 \vee \boldsymbol{v}_2 = (1, 4)$, and $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2) = (0, 0, 1)$. Therefore $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2) \neq \boldsymbol{u}_1 \wedge \boldsymbol{u}_2$, so that the profile $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ is not implementable, even though the implementable profile $\boldsymbol{v}_1$ implements $\boldsymbol{u}_1$ and the implementable profile $\boldsymbol{v}_2$ implements $\boldsymbol{u}_2$. ■

On the basis of the preceding discussion and Example 1, we may conclude that in general, $\boldsymbol{I}(X)$ is not a sublattice of $\boldsymbol{B}(X)$ and (for analogous reasons) $\boldsymbol{I}(Y)$ is not a sublattice of $\boldsymbol{B}(Y)$. We do know, though, that the inequality $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2) \leq \boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ must hold whenever $\boldsymbol{v}_1 \in \boldsymbol{I}(Y)$ implements $\boldsymbol{u}_1$ and $\boldsymbol{v}_2 \in \boldsymbol{I}(Y)$ implements $\boldsymbol{u}_2$: The implementation condition (4) implies that $\phi(x, y, \boldsymbol{v}_1(y)) \leq \boldsymbol{u}_1(x)$ and $\phi(x, y, \boldsymbol{v}_2(y)) \leq \boldsymbol{u}_2(x)$ hold for all $(x, y) \in X \times Y$. As type x obtains utility $\min\{\phi(x, y, \boldsymbol{v}_1(y)), \phi(x, y, \boldsymbol{v}_2(y))\}$ from choosing y when faced with the tariff $\boldsymbol{v}_1 \vee \boldsymbol{v}_2$, we thus obtain $\phi(x, y, \boldsymbol{v}_1(y) \vee \boldsymbol{v}_2(y)) \leq \boldsymbol{u}_1(x) \wedge \boldsymbol{u}_2(x)$ for all x and y and therefore $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2) \leq \boldsymbol{u}_1 \wedge \boldsymbol{u}_2$.

In the argument from the previous paragraph, the inequalities $\phi(x, y, \boldsymbol{v}_1(y)) \leq \boldsymbol{u}_1(x)$ and $\phi(x, y, \boldsymbol{v}_2(y)) \leq \boldsymbol{u}_2(x)$ both turn into equalities if $y \in Y_{\boldsymbol{v}_1}(x) \cap Y_{\boldsymbol{v}_2}(x)$ holds. For such y we thus have $\phi(x, y, \boldsymbol{v}_1(y) \vee \boldsymbol{v}_2(y)) = \boldsymbol{u}_1(x) \wedge \boldsymbol{u}_2(x)$, so that $\Phi(\boldsymbol{v}_1 \vee \boldsymbol{v}_2)$ attains its upper bound $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ if there exists a profile $\boldsymbol{y}$ satisfying $\boldsymbol{y}(x) \in Y_{\boldsymbol{v}_1}(x) \cap Y_{\boldsymbol{v}_2}(x)$ for all $x \in X$, meaning that both $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ implement $\boldsymbol{y}$. Recalling the definition of the set $\mathcal{U}_{\boldsymbol{y}}$ in equation (20) this proves the first statement in the following lemma. The proof for the second statement is analogous.

Lemma 3. *Let Assumption 1 hold. The set $\mathcal{U}_{\boldsymbol{y}}$ is a sublattice of $\boldsymbol{B}(X)$ for all $\boldsymbol{y} \in Y^X$ and the set $\mathcal{V}_{\boldsymbol{x}}$ is a sublattice of $\boldsymbol{B}(Y)$ for all $\boldsymbol{x} \in X^Y$.*

We will provide more substantial lattice results when considering the matching model in Section 4. In preparation for these, we use Lemma 3 to derive one more preliminary result.

Consider the set of profiles $\boldsymbol{u}$ that are both compatible with a given implementable assignment $\boldsymbol{y}$ and satisfy a participation constraint, that is, consider the set $\{\boldsymbol{u} \in \mathcal{U}_{\boldsymbol{y}} \mid \boldsymbol{u} \geq \underline{\boldsymbol{u}}\}$ for some profile $\underline{\boldsymbol{u}} \in \boldsymbol{B}(X)$ that we take to be continuous. As the intersection of the sublattice $\mathcal{U}_{\boldsymbol{y}}$ (Lemma 3) and the sublattice of profiles satisfying $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$, this set is also a sublattice. In the quasilinear case it is not difficult to see that this sublattice (i) is non-empty and (ii) has a minimum element, say $\boldsymbol{u}^*$, for which the participation constraint is binding, that is, $\boldsymbol{u}^*(x) = \underline{\boldsymbol{u}}(x)$ holds for some $x \in X$. The proof of the following result (in Appendix A.7) shows that these two additional properties do not require quasilinearity but hold under the weaker condition that the assignment under consideration is strongly implementable.

Lemma 4. *Let Assumption 1 hold and let $\underline{\boldsymbol{u}} \in \boldsymbol{C}(X)$ and $\underline{\boldsymbol{v}} \in \boldsymbol{C}(Y)$.*

[4.1] If $\boldsymbol{y} \in Y^X$ is strongly implementable, then the sublattice $\{\boldsymbol{u} \in \mathcal{U}_{\boldsymbol{y}} \mid \boldsymbol{u} \geq \underline{\boldsymbol{u}}\}$ has a minimum element $\boldsymbol{u}^*$ and this minimum element satisfies $\boldsymbol{u}^*(x) = \underline{\boldsymbol{u}}(x)$ for some $x \in X$.

[4.2] If $\boldsymbol{x} \in X^Y$ is strongly implementable, then the sublattice $\{\boldsymbol{v} \in \mathcal{V}_{\boldsymbol{x}} \mid \boldsymbol{v} \geq \underline{\boldsymbol{v}}\}$ has a minimum element $\boldsymbol{v}^*$ and this minimum element satisfies $\boldsymbol{v}^*(y) = \underline{\boldsymbol{v}}(y)$ for some $y \in Y$.

The main difficulty in establishing Lemma 4.1 (the other case is analogous) is to exclude the possibility that the minimum element $\boldsymbol{u}^*$ is strictly greater than $\underline{\boldsymbol{u}}$ for all $x \in X$. We resolve this difficulty by exploiting the lattice structure observed in Lemma 3 and the assumption of strong implementability to construct an increasing sequence of profiles in $\mathcal{U}_{\boldsymbol{y}}$ that satisfy $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$ for some $x \in X$ (but may violate the participation constraint) and then show (using Corollary 4) that this sequence has a limit that satisfies the participation constraint for all $x \in X$ and satisfies it with equality for some $x \in X$.[19]

4 Stability in Matching Models

This section shows how the results developed in Section 3 can be used to study stable outcomes in (two-sided) matching models. Section 4.1 introduces the matching model and defines the stability notions—stable outcomes and pairwise stable outcomes—that we consider. The notion of a pairwise stable outcome, which abstracts from participation constraints, is important because such outcomes can be characterized in terms of a pair of profiles implementing each other together with the argmax correspondences associated with these profiles. Section 4.2 develops this link. Section 4.3 then exploits it to show how familiar results for the existence of stable outcomes in matching models with a finite number of agents can be combined with our duality results to obtain, via a limiting argument, the existence of stable outcomes in matching models with an infinity of types. The role of the implementation duality in this argument is analogous to the role of (generalized) conjugate duality in McCann's proof (McCann, 1995) of the Kantorovich duality for optimal transport problems (see also Villani, 2009, Chapter 5).[20]

[19] In the quasilinear case a much simpler argument will do: Suppose $\boldsymbol{u}^*(x) > \underline{\boldsymbol{u}}(x)$ holds for all $x \in X$. As $\underline{\boldsymbol{u}}$ has been assumed to be continuous, $\boldsymbol{u}^*$ is continuous by Proposition 2, and X is compact, there then exists $\epsilon > 0$ such that $\boldsymbol{u}^*(x) - \epsilon \geq \underline{\boldsymbol{u}}(x)$ holds for all $x \in X$. In the quasilinear case the profile given by $\boldsymbol{u}^*(x) - \epsilon$ is an element of $\mathcal{U}_{\boldsymbol{y}}$, contradicting the minimality of $\boldsymbol{u}^*$.

[20] Previously, Gretsky, Ostroy, and Zame (1992) have used tools from optimal transport to establish existence of stable outcomes in matching models with perfectly transferable (quasilinear) utility. Kaneko and Wooders (1986, 1996) establish an existence result for a class of infinite cooperative games which includes

The main result in Section 4.4 is Proposition 8, which establishes that the sets of stable profiles are complete sublattices of the set of stable profile, thereby generalizing a corresponding result for matching models with a finite number of agents (Demange and Gale, 1985) to models with an infinity of types.

4.1 The Matching Model

To obtain a matching model, we add to our basic ingredients (X, Y, ϕ) a pair of finite non-zero Borel measures μ on X and ν on Y, describing the distribution of agent types on each side of the market, and a pair of continuous reservation utility profiles $\underline{\boldsymbol{u}} : X \to \mathbb{R}$ and $\underline{\boldsymbol{v}} : Y \to \mathbb{R}$, describing the utilities agents achieve when remaining unmatched. A matching model is then a collection $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$.

4.1.1 Matches and Outcomes

We follow the optimal transportation literature (Villani, 2009; Galichon, 2016) and Gretsky, Ostroy, and Zame (1992) in using a measure λ on $X \times Y$ to describe who is matched with whom and who remains unmatched. Formally, a match for a matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ is a Borel measure λ on $X \times Y$ satisfying the matching conditions

$$\lambda_X(\tilde{X}) := \lambda(\tilde{X} \times Y) \leq \mu(\tilde{X}) \tag{22}$$

$$\lambda_Y(\tilde{Y}) := \lambda(X \times \tilde{Y}) \leq \nu(\tilde{Y}) \tag{23}$$

for all measurable $\tilde{X} \subseteq X$ and $\tilde{Y} \subseteq Y$. We interpret $\lambda(\tilde{X} \times \tilde{Y})$ as identifying the mass of buyers from $\tilde{X}$ who are matched with sellers from $\tilde{Y}$. Condition (22) indicates that the mass of buyers with types in $\tilde{X}$, given by the marginal measure $\lambda_X(\tilde{X})$, who are matched to *some* seller cannot exceed the mass of these buyers, with mass $\mu(\tilde{X}) - \lambda_X(\tilde{X}) \geq 0$ of the agents in the set $\tilde{X}$ remaining unmatched. The interpretation of condition (23) is analogous.

An outcome is a triple $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ consisting of a match λ and a pair of utility profiles $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$ satisfying the (dual) feasibility conditions

$$\boldsymbol{u}(x) = \phi(x, y, \boldsymbol{v}(y)) \quad \text{and} \quad \boldsymbol{v}(y) = \psi(y, x, \boldsymbol{u}(x)) \quad \forall (x, y) \in \operatorname{supp}(\lambda) \tag{24}$$

for matched agents and the feasibility conditions

$$\boldsymbol{u}(x) \;=\; \underline{\boldsymbol{u}}(x) \quad \forall x \in \operatorname{supp}(\mu - \lambda_X) \tag{25}$$

$$\boldsymbol{v}(y) \;=\; \underline{\boldsymbol{v}}(y) \quad \forall y \in \operatorname{supp}(\nu - \lambda_Y). \tag{26}$$

for unmatched agents.[21] These feasibility conditions require that matched pairs receive utilities that can be generated in their matches and unmatched agents obtain their reservation

matching models with both perfectly and imperfectly transferable utility as special cases, but to do so resort to a notion of approximate feasibility. In work contemporaneous to ours, Greinecker and Kah (2016) obtain existence of stable outcomes for a broad class of matching problems (including problems with nontransferable utility) with an infinity of types, using tools quite different from the ones we employ.

[21] By specifying an outcome in terms of utility profiles we are imposing the equal treatment property that all agents of the same type receive the same utility level. Greinecker and Kah (2016) demonstrate that this is an innocent simplification. Similarly, by requiring equalities in (24) we are imposing efficiency within each match rather than obtaining this as an implication of stability.

utilities. Observe that we require feasibility for all types in the supports of μ and ν. This is in contrast to the approximate feasibility notion employed in Kaneko and Wooders (1986, 1996).

4.1.2 Stable Outcomes

An outcome for a matching model is stable if it satisfies the participation constraints

$$\begin{aligned} \boldsymbol{u}(x) &\geq \underline{\boldsymbol{u}}(x) \quad \forall x \in \operatorname{supp}(\nu) &(27)\\ \boldsymbol{v}(y) &\geq \underline{\boldsymbol{v}}(y) \quad \forall y \in \operatorname{supp}(\mu) &(28)\end{aligned}$$

and the (dual) incentive constraints

$$\boldsymbol{u}(x) \geq \phi(x, y, \boldsymbol{v}(y)) \quad \text{and} \quad \boldsymbol{v}(y) \geq \psi(y, x, \boldsymbol{u}(x)) \quad \forall (x, y) \in \operatorname{supp}(\nu) \times \operatorname{supp}(\mu). \tag{29}$$

A match or profile will be called stable if it is part of a stable outcome.

The stability conditions require that, as indicated by (27)–(28), no matched agent in the support of one of the type distributions would rather be unmatched, and, as indicated by (29), no pair of agents in the supports of the type distributions can achieve strictly higher utilities by matching with each other than by sticking to the outcome under consideration.

Note that conditions (24)–(29) impose no constraints whatsoever on types that do not appear in the supports of the type distribution and, further, (29) does not preclude the possibility that some type x in the support of μ might prefer to match with a type outside of the support of ν (and vice versa). In essence, we are thus treating types that lie outside the support of the type-distributions as being non-existent in the definition of stable outcomes.[22]

Remark 7 (Stable Outcomes in Finite-Support Matching Models)**.** To extend results from matching models with a finite number of agents to matching models with an infinite number of types we consider *finite-support matching models*. Formally, we say that the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ has finite support if there exists $(x_1, \ldots, x_m) \in X^m$ and $(y_1, \ldots, y_n) \in Y^n$ such that the measures μ and ν on X and Y satisfy

$$\mu(\tilde{X}) = \sum_{i=1}^{m} \delta_{x_i}(\tilde{X}) \text{ and } \nu(\tilde{Y}) = \sum_{j=1}^{n} \delta_{y_i}(\tilde{Y})$$

for all measurable $\tilde{X} \subseteq X$ and measurable $\tilde{Y} \subseteq Y$, where m and n are natural numbers and δ_x (and similarly δ_y) is the Dirac measure on X assigning mass 1 to x. The $(x_1, \ldots, x_m)$ (and similarly $(y_1, \ldots, y_n)$) need not be distinct, so that there could be at most m distinct types of buyers and n distinct types of sellers, though there could be many buyers (or sellers) of a single type and hence fewer distinct types.

We can interpret a finite-support matching model as a familiar *matching model with a finite number of agents*, where the latter includes a finite set of buyers $I = \{1, \ldots, m\}$

[22] One may then wonder why we do not simply exclude such types from the model by assuming that μ and ν have full support. The answer is that it will prove technically convenient to consider models—like the finite-support matching models introduced in the following Remark 7—which violate such a full support condition.

and a finite set of sellers $J = \{1, \ldots, n\}$. Buyer i has type x_i and seller j has type y_j. The standard definition of a match for such a model (see, for instance, Roth and Sotomayor (1990, Definition 9.1)) is equivalent to specifying a measure ρ on $I \times J$ that satisfies $\rho(i,j) \in \{0,1\}$ for all $(i,j) \in I \times J$, $\sum_{j \in J} \rho(i,j) \leq 1$ for all $i \in I$, and $\sum_{i=I} \rho(i,j) \leq 1$ for all $j \in J$. A stable outcome then consists of such a match and a specification of utility profiles $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ satisfying the natural counterparts to our feasibility and stability conditions (e.g. (24) becomes $u_i = \phi(x_i, y_j, v_j)$ for all (i,j) satisfying $\rho(i,j) = 1$ and (29) becomes $u_i \geq \phi(x_i, y_j, v_j)$ for all $(i,j) \in I \times J$).

It is well-known that stable outcomes for a matching model with a finite number of agents exist if the characteristic function describing the utility frontier available to a pair of matched agents satisfies our Assumption 1 (Roth and Sotomayor, 1990, Section 9.4). Further, every such stable outcome satisfies the equal treatment property (i.e., $x_i = x_{i'}$ implies $u_i = u_{i'}$ and $y_j = y_{j'}$ implies $v_j = v_{j'}$). This allows us to map any stable outcome $(\rho, u_1, \ldots, u_m, v_1, \ldots, v_n)$ for the matching model with a finite number of agents into a stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ for our finite-support matching model (see Appendix A.8 for the details of this construction). In particular, every finite-support matching model has a stable outcome. ■

4.1.3 Pairwise Stable Outcomes in Balanced Matching Models

We say that a matching model is balanced if $\mu(X) = \nu(Y)$ holds, so that the mass of buyers and sellers are identical. A match λ for a balanced matching model is full if the inequalities in (22) and (23) hold as equalities,

$$\lambda(\tilde{X} \times Y) = \mu(\tilde{X}) \tag{30}$$

$$\lambda(X \times \tilde{Y}) = \nu(\tilde{Y}) \tag{31}$$

for all measurable $\tilde{X} \subseteq X$ and $\tilde{Y} \subseteq Y$, indicating that there are no unmatched agents. An outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ for a balanced matching model is full if it features a full match. Observe that for any full match the feasibility conditions (25) and (26) are vacuous (because $\mathrm{supp}(\mu - \lambda_X) = \mathrm{supp}(\nu - \lambda_Y) = \emptyset$), so that an outcome is full if and only if it satisfies (24), (30), and (31). In line with our definition of profiles $\boldsymbol{u}$ or $\boldsymbol{v}$ satisfying an initial condition (cf. Section 2.4), we say that a full outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ for a balanced matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfies initial condition $(x_0, u_0) \in X \times \mathbb{R}$ if $\boldsymbol{u}(x_0) = u_0$, and satisfies initial condition $(y_0, v_0) \in Y \times \mathbb{R}$ if $\boldsymbol{v}(y_0) = v_0$.

We define a full outcome to be pairwise stable if it satisfies the incentive constraints (29). Note that full matches and full outcomes exist only for balanced matching models and that whenever we call an outcome, match, or profile pairwise stable, it is implied that it is part of a full outcome. Note, further, that whether or not an outcome is pairwise stable is independent of the specification of the reservation utility profiles $\underline{\boldsymbol{u}}$ and $\underline{\boldsymbol{v}}$. On the other hand, a pairwise stable outcome is stable if and only if it satisfies the participation constraints (27) and (28).

Our definition of a full match for a balanced matching model is identical to the definition of a transportation (or transference) plan in the literature on optimal transport. This allows us to borrow results from this literature when analysing full matches and full outcomes. For instance, it is well-known that (under our maintained assumption that X and Y are compact) the set of full matches is compact in the topology of weak convergence of measures (cf. Villani, 2009, p. 45).

4.1.4 Deterministic Matches

In many economic applications it is natural to focus on full matches that can be described in terms of assignments, thereby identifying for all agent types on one side of the matching market a unique partner on the other side with whom they are matched. This is captured by the notion of a deterministic match—corresponding to the notion of a deterministic coupling or transport map in the optimal transportation literature (Villani, 2009, p.6)—defined in the following.[23]

We say that a measure λ on the set $X \times Y$ is deterministic and denote it by $\lambda_{\boldsymbol{y}}$ if there exists a measurable assignment $\boldsymbol{y}$ such that

$$\lambda(\tilde{X} \times \tilde{Y}) = \mu(\{x \in \tilde{X} | \boldsymbol{y}(x) \in \tilde{Y}\}) \tag{32}$$

for measurable $\tilde{X} \subseteq X$ and $\tilde{Y} \subseteq Y$. If such a deterministic measure λ is a full match in the balanced matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$, then it is a deterministic match.

If $\lambda_{\boldsymbol{y}}$ is a deterministic match then the assignment $\boldsymbol{y}$ must be measure preserving (and hence necessarily measurable), i.e., $\nu(\tilde{Y}) = \mu(\boldsymbol{y}^{-1}(\tilde{Y}))$ must hold for all measurable $\tilde{Y} \subseteq Y$.[24]

4.2 Connecting Implementability and Pairwise Stability

With a quasilinear generating function $\phi(x, y, v) = f(x, y) - v$ a match is pairwise stable if and only if it maximizes the surplus $\int_{X \times Y} f(x, y) d\lambda(x, y)$ over the set of full matches. Standard results from the optimal transport literature then imply that a full match λ is pairwise stable if and only if it its support is contained in $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ for a pair of profiles $(\boldsymbol{u}, \boldsymbol{v})$ implementing each other, and that for such a pair of profiles the full outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a pairwise stable outcome (cf. Galichon, 2016, Chapters 6 and 7). These results carry over without any changes to the general case:

Proposition 5. *Let Assumptions 1 hold and let the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be balanced.*

[5.1] If λ is a full match, then $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a full outcome if and only if $\text{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$.

[5.2] If $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a full outcome and (i) $\boldsymbol{u}$ implements $\boldsymbol{v}$ or (ii) $\boldsymbol{v}$ implements $\boldsymbol{u}$, then $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is pairwise stable.

[23]We focus on assignments $\boldsymbol{y} \in Y^X$ with all our definitions and observations carrying over to assignments $\boldsymbol{x} \in X^Y$ in the obvious way.

[24]In general, pairwise stable deterministic matches do not exist in balanced matching models, even when the existence of measure-preserving assignments is assured (e.g. when μ is atomless) and the generating function is quasilinear. Villani (2009, Example 4.9) provides a simple example for an optimal-transport problem (with both μ and ν atomless) which has no deterministic solution. This example is easily modified to demonstrate the non-existence of pairwise stable deterministic matches. See also Gretsky, Ostroy, and Zame (1992) for an extended discussion of related existence questions in the context of a two-sided matching model and an argument which, when transferred to our setting, suggests that it is possible to interpret any of the full matches we consider as measure-preserving bijections between suitably enlarged measure spaces. Greinecker and Kah (2016) pursue such a construction.

[5.3] If $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a pairwise stable outcome, then there exists profiles $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ with the properties that (i) $\tilde{\boldsymbol{u}}(x) = \boldsymbol{u}(x)$ on the support of μ and $\tilde{\boldsymbol{v}}(y) = \boldsymbol{v}(y)$ on the support of ν, (ii) $(\lambda, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}})$ is a pairwise stable outcome for $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$, and (iii) $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ implement each other.

Proof. [5.1] If λ is a full match, then (24) is necessary and sufficient for $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ to be a full outcome. By definition of $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ (see (16)), condition (24) holds if and only if $\text{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$.

[5.2] If $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a full outcome, then (24) and (30) and (31) hold. Therefore, (29), which holds if $\boldsymbol{v}$ implements $\boldsymbol{u}$ or $\boldsymbol{v}$ implements $\boldsymbol{u}$, is sufficient for $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ to be pairwise stable.

[5.3] See Appendix A.9 ■

If the type measures μ and ν both have full support, Proposition 5.3 is also immediate from the definitions, and is the statement that if $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is pairwise stable, then $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other. Otherwise Proposition 5.3 indicates that the profiles $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ in any pairwise stable outcome can be adjusted outside the supports of μ and ν in such a way that the suitably adjusted profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other. In either case, in conjunction with the other two parts of the proposition, we obtain the conclusion that a full match λ is pairwise stable if and only if it satisfies $\text{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ for a pair of profiles implementing each other.

For a deterministic match $\lambda_{\boldsymbol{y}}$ with implementable $\boldsymbol{y}$, it is almost immediate from Propositions 4 and 5 that $\lambda_{\boldsymbol{y}}$ is a pairwise stable match. Obtaining a converse statement is more difficult because $\text{supp}(\lambda_{\boldsymbol{y}}) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ does not necessarily imply that the graph of $\boldsymbol{y}$ is contained in $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$. Appendix A.10 proves:

Lemma 5. *Let Assumption 1 hold, let the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be balanced, and let λ be a deterministic match. Then λ is a pairwise stable match if and only if there exists an implementable $\boldsymbol{y} \in Y^X$ such that $\lambda = \lambda_{\boldsymbol{y}}$ holds.*

4.3 Existence of (Pairwise) Stable Outcomes

We begin by exploiting our duality results to establish the existence of pairwise stable outcomes in balanced matching models satisfying arbitrary initial conditions. Appendix A.11 proves:

Proposition 6. *Let Assumptions 1 hold and let the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be balanced. Then for every initial condition (y_0, v_0) (and similarly for every initial condition (x_0, u_0)), there exists a pairwise stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ satisfying the initial condition (y_0, v_0) in which $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other.*

The proof of Proposition 6 begins by considering balanced finite-support matching models (cf. Remark 7) with at most n (distinct) types of buyers and at most n (distinct) types of sellers, and exploiting Lemma 5 in Demange and Gale (1985) to show that such a finite-support matching model has a pairwise stable outcome $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ satisfying the required initial condition. In addition, Proposition 5.3 ensures that we can take the profiles $(\boldsymbol{u}_n, \boldsymbol{v}_n)$ to implement each other. We next construct a sequence of such finite-support balanced matching models for which the associated measures μ_n and ν_n converge weakly to the target measures μ and ν. Prokhorov's theorem allows us to conclude that the sequence of measures $(\lambda_n)_{n=1}^{\infty}$ has a subsequence converging weakly to a full match λ^*. Exploiting the fact that the initial condition holds along the sequences to show that the sequences of profiles $(\boldsymbol{u}_n)_{n=1}^{\infty}$ and $(\boldsymbol{v}_n)_{n=1}^{\infty}$ are bounded, it becomes a straightforward consequence of our duality results that these sequences have subsequences converging to profiles $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$ implementing each other and that, further, $(\lambda^*, \boldsymbol{u}^*, \boldsymbol{v}^*)$ is a pairwise stable outcome for $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfying the initial condition. This gives us the desired result.

To go from the existence result for pairwise stable outcomes in balanced matching models in Proposition 6 to an existence result for stable outcomes in any matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfying Assumption 1, we consider an augmented matching model. As in a similar construction in Chiappori, McCann, and Nesheim (2010), in this augmented model the type spaces differ from X and Y by the addition of dummy types x_0 and y_0 on each side of the market. Adding the dummy types x_0 and y_0 transforms the original matching model into a balanced matching model in which (i) being unmatched in the original model corresponds to being matched with a dummy agent in the augmented matching model, (ii) for an appropriate choice of initial conditions, a pairwise stable outcome in the augmented model corresponds to a stable outcome in the original model, and (iii) Assumption 1 holds for the augmented model. Given these properties of the augmented matching model, Proposition 6 implies the existence of a stable outcome for the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. The proof of the following result, in Appendix A.12, shows how to construct an augmented matching model with the requisite properties.

Corollary 5. *Let Assumption 1 hold. There exists a stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ for the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$.*

4.4 Lattice Structure of (Pairwise) Stable Profiles

The main result of this section is Proposition 8, which establishes that the sets of stable profiles are complete sublattices of the sets of bounded profiles. As in Section 4.3, we first establish a preliminary result for pairwise stable outcomes. This again reflects the logical structure of the arguments, with the more substantive work occurring in the first step and the second following from the link between pairwise stable full outcomes and stable outcomes which underlies the proof of Corollary 5. The results for pairwise stable outcomes will also be of independent use when we turn to the principal-agent model.

The following assumption simplifies the exposition by ensuring (from Proposition 5.3) that in every pairwise stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$, the profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other.[25]

Assumption 2. *The type measures μ and ν have full support.*

[25] This allows us to avoid the adjustments to $\boldsymbol{u}$ or $\boldsymbol{v}$ outside the supports of μ and ν that appear in the proof of Proposition 5.3.

4.4.1 The Lattice of Pairwise Stable Profiles

Let

$$\mathbb{U} = \{\boldsymbol{u} \in \boldsymbol{B}(X) \mid (\lambda, \boldsymbol{u}, \boldsymbol{v}) \text{ is pairwise stable for some full match } \lambda \text{ and } \boldsymbol{v} \in \boldsymbol{B}(Y)\}$$
$$\mathbb{V} = \{\boldsymbol{v} \in \boldsymbol{B}(Y) \mid (\lambda, \boldsymbol{u}, \boldsymbol{v}) \text{ is pairwise stable for some full match } \lambda \text{ and } \boldsymbol{u} \in \boldsymbol{B}(X)\}$$

denote the sets of pairwise stable profiles in a balanced matching model. From Proposition 6 the sets $\mathbb{U}$ and $\mathbb{V}$ are non-empty if Assumption 1 holds. The following result shows that they are also closed sublattices (of $\boldsymbol{B}(X)$, resp. of $\boldsymbol{B}(Y)$).

Proposition 7. *Let Assumptions 1 and 2 hold and let the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be balanced. The sets $\mathbb{U}$ and $\mathbb{V}$ of pairwise stable profiles are closed sublattices.*

Appendix A.13 contains the proof. The idea behind the proof that $\mathbb{U}$ and $\mathbb{V}$ are sublattices is the same as the one behind the Decomposition Lemma in Demange and Gale (1985): Given two pairwise stable outcomes $(\lambda_1, \boldsymbol{u}_1, \boldsymbol{v}_1)$ and $(\lambda_2, \boldsymbol{u}_2, \boldsymbol{v}_2)$ we show that both X and Y can be partitioned into two sets each, say X into X_1 and X_2 and Y into Y_1 and Y_2, such that both λ_1 and λ_2 only match buyer types from X_1 with seller types in Y_1 and buyer types in X_2 with seller types in Y_2. Further, when faced with $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$, all buyers in X_1 prefer to be matched as under λ_1, whereas the reverse preference holds for buyers in X_2. Constructing a measure λ_3 on $X \times Y$ by matching the types in X_1 and Y_1 as under λ_1 and the types in X_2 and Y_2 as under λ_2 then yields a pairwise stable outcome $(\lambda_3, \boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2)$. An analogous argument may be used to obtain the existence of a full match λ_4 such that $(\lambda_4, \boldsymbol{u}_1 \wedge \boldsymbol{u}_2, \boldsymbol{v}_1 \vee \boldsymbol{v}_2)$ is a pairwise stable outcome. Taken together, the existence of the pairwise stable outcomes $(\lambda_3, \boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2)$ and $(\lambda_4, \boldsymbol{u}_1 \wedge \boldsymbol{u}_2, \boldsymbol{v}_1 \vee \boldsymbol{v}_2)$ implies that both $\mathbb{U}$ and $\mathbb{V}$ are sublattices. The closedness claim in the statement of the proposition follows from the same arguments we have used in the proof of Proposition 6 to establish that the limit of the pairwise stable outcomes in the approximating finite-support matching models considered there is pairwise stable.

The proof of Proposition 7 would be much simpler if we could assume that all pairs $(\boldsymbol{u}_1, \boldsymbol{v}_1)$ and $(\boldsymbol{u}_2, \boldsymbol{v}_2)$ of stable profiles are compatible with the *same* stable match λ.[26] In that case an argument analogous to that of Lemma 3 would yield that $\mathbb{U}$ and $\mathbb{V}$ are sublattices. However, as illustrated by Roth and Sotomayor (1990, Example 9.6, p. 225) and Quint (1994, Example 6.1, p. 612), this is generally not the case if the generating function is not quasilinear.

Recall that Lemma 4 in Section 3.4.2 has established that the set of profiles $\mathcal{U}_{\boldsymbol{y}}$ compatible with a *given* strongly implementable assignment $\boldsymbol{y}$ satisfying a participation constraint has a minimum element in which the participation constraint is binding for some type. The essential properties of $\mathcal{U}_{\boldsymbol{y}}$ used in this proof were that the set $\mathcal{U}_{\boldsymbol{y}}$ is a closed (Corollary 4.1) sublattice (Lemma 3) of implementable profiles containing an element for every possible

[26]This is trivially true if there is a unique stable match, as is the case under a strict single crossing condition (Proposition 12 in Section 6). It is also true with a quasilinear generating function, as with transferable utility all stable profiles are compatible with the same stable match; see Roth and Sotomayor (1990, Corollary 8.7, p. 207) for finite matching models and Gretsky, Ostroy, and Zame (1999), who also use this fact to establish a counterpart to our Proposition 8 below (Gretsky, Ostroy, and Zame, 1999, Proposition 5), for a model with an infinity of types.

initial condition (by strong implementability). The set of pairwise stable profiles $\mathbb{U}$ satisfies the same properties: it is a closed (Proposition 6) sublattice (Proposition 7) of implementable profiles (Proposition 5.2) with the set $\{u \in \mathbb{U} \mid \boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)\}$ non-empty for all $x \in X$ (Proposition 6). Therefore, the following counterpart to Lemma 4 holds for the sets of pairwise stable profiles (with the proof being identical):

Corollary 6. *Let Assumptions 1 and 2 hold and let $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be a balanced matching model. Then the set of pairwise stable buyer profiles satisfying the participation constraint $\boldsymbol{u}(x) \geq \underline{\boldsymbol{u}}(x)$ for all $x \in X$ has a minimum element $\boldsymbol{u}^*$, at which the equality $\boldsymbol{u}^*(x) = \underline{\boldsymbol{u}}(x)$ holds for some $x \in X$. Similarly, the set of pairwise stable seller profiles satisfying the participation constraints $\boldsymbol{v}(y) \geq \underline{\boldsymbol{v}}(y)$ for all $y \in Y$ has a minimum element v^*, at which the equality $\boldsymbol{v}^*(y) = \underline{\boldsymbol{v}}(y)$ holds for some $y \in Y$.*

4.4.2 The Lattice of Stable Profiles

The connection between pairwise stability in balanced matching models and stability in arbitrary matching models underlying the proof of Corollary 5 in Section 4.3 allows us to extend our results about the lattice structure of pairwise stable profiles to results about the lattice structure of stable profiles.

First, we use Proposition 7 to show that the sets of stable buyer and seller profiles are complete sublattices. Appendix A.14 proves:

Proposition 8. *Let Assumptions 1 and 2 hold. The sets of stable seller profiles and stable buyer profiles of the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ are complete sublattices.*

Second, we use Corollary 6 to establish a counterpart to Lemma 3 in Demange and Gale (1985), asserting that in a balanced matching model *both* the minimum buyer stable profile u^* and the minimum seller stable profile v^* feature binding participation constraints.[27]

Corollary 7. *Let Assumptions 1 and 2 hold and let $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be a balanced matching model. Then the minimum stable buyer profile $\boldsymbol{u}^*$ satisfies $\boldsymbol{u}^*(x) = \underline{\boldsymbol{u}}(x)$ for some $x \in X$ and the minimum stable seller profile $\boldsymbol{v}^*$ satisfies $\boldsymbol{v}^*(y) = \underline{\boldsymbol{v}}(y)$ for some $y \in Y$.*

Proof. The claim is immediate from the feasibility conditions (25)–(26) unless all stable outcomes are fully matched. We therefore suppose this to be the case. The set of stable outcomes then coincides with the set of pairwise stable outcomes $(\lambda, \boldsymbol{u}, \boldsymbol{v})$, satisfying the participation constraints $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ and $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$. Recalling that for any pairwise stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ the profiles $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other (Proposition 5.3), the result then follows from Corollary 6 provided that the profiles $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$ appearing in the statement of that corollary satisfy $\Psi \boldsymbol{u}^* \geq \underline{\boldsymbol{v}}$ and $\Phi \boldsymbol{v}^* \geq \underline{\boldsymbol{u}}$. Recalling that the implementation maps are order reserving, these conditions must be satisfied (as otherwise the set of stable profiles would be empty). ■

[27] In an unbalanced matching model (satisfying $\mu(X) \neq \nu(Y)$) it is trivially the case that in every outcome there are unmatched agents on the "long side" of the market. By the feasibility conditions (25)–(26) such unmatched agents receive their reservation utility, so that either the minimum buyer stable profile $\boldsymbol{u}^*$ or the minimum seller stable profile $\boldsymbol{v}^*$ features a binding participation constraint. In particular, if $\mu(X) > \nu(Y)$, then there exists $x \in X$ satisfying $\boldsymbol{u}^*(x) = \underline{\boldsymbol{u}}(x)$ and, similarly, if $\mu(X) < \nu(Y)$, then there exists $y \in Y$ satisfying $\boldsymbol{v}^*(y) = \underline{\boldsymbol{v}}(y)$. Note the existence of $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$ is ensured because the sets of stable profiles are complete sublattices (Proposition 8).

5 Optimal Outcomes in Principal-Agent Models

This section applies our characterization of implementable profiles and assignments to adverse-selection principal-agent models. Section 5.1 formulates the principal's problem as choosing a measure λ on $X \times Y$, as well as a rent function $\boldsymbol{u}$ and a tariff $\boldsymbol{v}$, subject to incentive and participation constraints. Remark 8 explains how this formulation allows us to interpret triples $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ that satisfy the incentive constraints in the principal's problem as pairwise stable outcomes in a balanced matching model.

Section 5.2 reformulates the principal's problem as a nonlinear pricing problem in which the maximization is over a set of tariffs, and then uses this reformulation to establish that the principal's problem has a solution. Moreover, it has a solution in which the measure λ chosen by the principal is deterministic and thus corresponds to the choice of an optimal assignment. Our duality results play a central role in this existence argument, with Corollary 3.1 ensuring that we can model the principal as choosing an *implementable* tariff, and Corollary 4 ensuring that the resulting feasible set is compact.[28]

Section 5.3 considers whether the agent's participation constraint must be binding in a solution to the principal's problem. It provides an example to show that this need not be the case and two conditions sufficient to ensure it, namely that either every implementable profile is strongly implementable or that the principal's utility function exhibits private values. The first result is consistent with our view of strong implementability as a useful generalization of quasilinearity, while the second makes essential use of the connections to the matching model.

Section 5.4 introduces the possibility of exclusion into the model.

5.1 The Principal-Agent Model

To obtain a principal-agent model, we add to our basic elements (X, Y, ϕ) a function $\pi : X \times Y \times \mathbb{R} \to \mathbb{R}$, describing the principal's utility of receiving payment v from agent type x who takes decision y, a Borel measure μ on the set X describing the distribution of agent types, and a continuous profile $\underline{\boldsymbol{u}} : X \to \mathbb{R}$ describing the agent's reservation utilities. A principal-agent model is then a collection $(X, Y, \phi, \mu, \pi, \underline{\boldsymbol{u}})$.

Assumption 3. *The function π is continuous, strictly increasing in its third argument, and satisfies $\pi(x, y, \mathbb{R}) = \mathbb{R}$ for all $(x, y) \in X \times Y$. The type measure μ has full support.*

Let $\mathbb{M}$ be the set of Borel measures on $X \times Y$ whose marginal distribution on the set X equals μ. We formulate the principal's problem as choosing a triple $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ consisting of a

[28]Obtaining compactness of the feasible set (and the requisite continuity properties of the principal's objective function) is the main difficulty in the existence proofs in Kahn (1993), Carlier (2001), and Carlier (2002), who consider special cases of the principal-agent model in which the agent's utility function is quasilinear. Using the special structure resulting from the imposition of a single crossing condition when X and Y are intervals, Jullien (2000) provides a straightforward existence argument which uses Helly's selection theorem in lieu of compactness arguments. Working without quasilinearity, the existence proofs in Page (1991, 1992, 1997) and Balder (1996) impose compactness as an assumption on the set of feasible contracts. Recently, allowing for stochastic contracts, Kadan, Reny, and Swinkels (2017) have obtained a very general existence result for principal-agent models with both adverse selection and moral hazard using tools rather different from the ones we employ. We explain in Section 7 how our approach can be extended to allow for stochastic contracts and moral hazard.

measure $\lambda \in \mathbb{M}$, a utility profile $\boldsymbol{u} \in \boldsymbol{B}(X)$, and a tariff $\boldsymbol{v} \in \boldsymbol{B}(Y)$ to maximize

$$\int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(x, y) \tag{33}$$

subject to the feasibility constraints

$$\begin{aligned} &\boldsymbol{v} \text{ implements } \boldsymbol{u} \\ &\operatorname{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}} \\ &\boldsymbol{u} \geq \underline{\boldsymbol{u}}. \end{aligned}$$

If λ is a deterministic measure $\lambda_{\boldsymbol{y}}$ (cf. (32)), then the first two constraints in this maximization problem are the standard incentive constraints, requiring that $\boldsymbol{u}$ is the rent function that results when each agent type maximizes against the tariff $\boldsymbol{v}$ and that all agent types x are assigned to one of their optimal decisions $\boldsymbol{y}(x) \in Y_{\boldsymbol{v}}(x)$. Intuitively, for measures $\lambda \in \mathbb{M}$ that are not deterministic, the second of these conditions is weakened to allow the principal to randomize over the set of decisions that are optimal for the agent.

The principal's expected utility in (33) is well-defined for any feasible $(\lambda, \boldsymbol{u}, \boldsymbol{v})$: Because $\operatorname{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$, we have (cf. the definition of $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ in (16)) $\boldsymbol{v}(y) = \psi(y, x, \boldsymbol{u}(x))$ for all $(x, y) \in \operatorname{supp}(\lambda)$, and hence

$$\int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(x, y) = \int_X \int_Y \pi(x, y, \psi(y, x, \boldsymbol{u}(x))) d\lambda(x, y), \tag{34}$$

where the latter integral is well-defined because π, ψ, and the implementable profile $\boldsymbol{u}$ are continuous (the latter by Proposition 2). A useful implication is that the principal's payoff can be written in terms of only the measure λ and rent function $\boldsymbol{u}$, implying that any two feasible outcomes $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ and $(\lambda, \boldsymbol{u}, \tilde{\boldsymbol{v}})$ give the same payoff to the principal.

Remark 8 (Pairwise Stability and Feasibility in the Principal's Problem)**.** Consider a triple $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ that satisfies the incentive constraints in the principal's problem, that is, $\boldsymbol{v}$ implements $\boldsymbol{u}$ and $\operatorname{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$. Define the measure ν on Y by setting $\nu(\tilde{Y}) = \lambda_Y(\tilde{Y})$ for all measurable $\tilde{Y} \subset Y$ and specify an arbitrary continuous reservation utility profile $\underline{\boldsymbol{v}}$. Then λ is a full match for the balanced matching problem $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. Further it is immediate from Proposition 5 that $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is pairwise stable in this matching problem. Vive versa, if $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is pairwise stable for a matching problem $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ in which μ has full support, then $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ satisfies the incentive constraints in any principal-agent model $(X, Y, \phi, \mu, \pi, \underline{\boldsymbol{u}})$ in which π has the properties from Assumption 3. See Carlier (2003, Theorem 2) and, more recently, Dworczak and Zhang (2017) for related observations in the quasilinear case. ■

5.2 Existence of a Solution to the Principal's Problem

To obtain our existence result, we begin by transforming the principal's problem into a nonlinear pricing problem over the set of implementable tariffs $\boldsymbol{v} \in \boldsymbol{I}(Y)$. Towards this end, define the function $F : \boldsymbol{I}(Y) \times \mathbb{M} \to \mathbb{R}$ by

$$F(\boldsymbol{v}, \lambda) = \int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(x, y) \tag{35}$$

and define the correspondence $G : \boldsymbol{I}(Y) \to \mathbb{M}$ by

$$G(\boldsymbol{v}) = \{\lambda \in \mathbb{M} : \operatorname{supp}(\lambda) \in \Gamma_{\Phi \boldsymbol{v}, \boldsymbol{v}}\}. \tag{36}$$

Also, for $\boldsymbol{v} \in \boldsymbol{I}(Y)$ let

$$\Pi(\boldsymbol{v}) = \max_{\lambda \in G(\boldsymbol{v})} F(\boldsymbol{v}, \lambda). \tag{37}$$

Observe that $F(\boldsymbol{v}, \lambda)$ is nothing but the objective function of the principal's problem specified in (33). The heuristic interpretation of (37) therefore is that $\Pi(\boldsymbol{v})$ specifies the maximal payoff the principal can obtain by probabilistically assigning agents to decision that are optimal for them when facing the implementable tariff $\boldsymbol{v}$ (which is captured by the constraint $\lambda \in G(v)$). Appendix A.15 shows that this problem has a solution for every implementable tariff, so that the function $\Pi : \boldsymbol{I}(Y) \to \mathbb{R}$ is well-defined. Further, it shows:

Lemma 6. *Let Assumptions 1 and 3 hold. The function $\Pi : \boldsymbol{I}(Y) \to \mathbb{R}$ is upper semicontinuous. If $\boldsymbol{v}^*$ solves*

$$\max_{\{\boldsymbol{v} \in \boldsymbol{I}(Y) : \boldsymbol{v} \leq \Psi \underline{\boldsymbol{u}}\}} \Pi(\boldsymbol{v}), \tag{38}$$

then there exists $\lambda^* \in G(v^*)$ such that the triple $(\lambda^*, \Phi \boldsymbol{v}^*, \boldsymbol{v}^*)$ solves the principal's problem.

The first step in the proof of Lemma 6 uses Corollary 3.1 to show that replacing an arbitrary tariff $\boldsymbol{v}$ in a feasible triple $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ with the implementable tariff $\Psi \boldsymbol{u}$ results in another feasible triple. Doing so leaves the principal's expected payoff unchanged (cf. (34)). This allows us to reduce the principal's problem to the choice of an implementable tariff $\boldsymbol{v}$ and an associated $\lambda \in G(\boldsymbol{v})$, with the agent's utility profile given by the rent function $\boldsymbol{u} = \Phi \boldsymbol{v}$. The continuity of implementable profiles $\boldsymbol{v}$ (Proposition 2) and the compactness of the set of measures $\mathbb{M}$ (Prokhorov's theorem (Shiryaev, 1996, p. 318)) then ensure that the function F and the correspondence G are sufficiently well-behaved to imply the upper semicontinuity of the function Π. Maximizing this function subject to the constraint that the associated rent function $\Phi \boldsymbol{v}$ satisfies the participation constraints $\Phi \boldsymbol{v} \geq \underline{\boldsymbol{u}}$, which we may rewrite as $\boldsymbol{v} \leq \Psi \underline{\boldsymbol{u}}$, then yields an optimal tariff $\boldsymbol{v}^*$ that together with the associated measure λ^*, and induced rent function $\boldsymbol{u}^* = \Phi \boldsymbol{v}^*$ solves the principal's problem.

To show the existence of a solution to the principal's problem it remains to show that the nonlinear pricing problem (38) in the statement of Lemma 6 has a solution. To do so, we begin by observing that the feasible set of the nonlinear pricing problem is bounded above by $\Psi \underline{\boldsymbol{u}}$. 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. We can thereby restrict the choice set in the nonlinear pricing problem to a closed and bounded set of tariffs. Moreover, and crucially, the maximization in (38) is over a set of implementable profiles, and we have established in Corollary 4.3 that closed and bounded sets of implementable profiles are compact. As Π is upper semicontinuous (Lemma 6), an application of Weierstrass' extreme value theorem then yields the existence of a solution to the nonlinear pricing problem. Appendix A.16 shows, in addition, that the measure in the associated solution to the principal's problem can be "purified" to obtain a solution to the principal's problem featuring a deterministic match:

Proposition 9. *Let Assumptions 1 and 3 hold. Then there exists a solution* $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ *to the principal's problem in which* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *implement each other and* λ *is deterministic.*

In light of Proposition 9, we hereafter restrict attention to solutions to the principal's problem in which $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other. When convenient, we will further restrict attention to deterministic solutions $(\lambda_{\boldsymbol{y}}, \boldsymbol{u}, \boldsymbol{v})$ to the principal's problem, sometimes denoting these by $(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$ to simplify notation.

5.3 Is the Participation Constraint Binding?

As the principal must respect the agent's participation constraint when choosing an optimal tariff, we have $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ in any solution to the principal's problem. Here we ask whether the agent's participation constraint must be binding in the sense that there exists some $x \in X$ satisfying $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$.[29]

If all implementable assignments $\boldsymbol{y}$ are strongly implementable, then the answer is straightforward from the lattice result in Lemma 4. Appendix A.17 shows:

Proposition 10. *Let Assumptions 1 and 3. If every implementable assignment* $\boldsymbol{y}$ *is strongly implementable, then the participation constraint is binding in any solution to the principal's problem.*

As we have noted in Section 2.4, quasilinearity of the agent's utility function is sufficient but not necessary for every implementable assignment to be strongly implementable. Proposition 10 establishes that it is not quasilinearity but this latter, weaker property which ensures that the participation constraint is binding in any solution to the principal's problem.

In the absence of strong implementability, the conclusion of Proposition 10 may fail:

Example 2. Let $X = \{1, 2\}$ and $Y = \{1, 2\}$. The generating function is given by

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

Let $\mu(1) = \mu(2) = 1/2$ and $\underline{\boldsymbol{u}}(1) = \underline{\boldsymbol{u}}(2) = 0$. Then Assumptions 1 and 3 hold for any specification of the principal's utility function π which is strictly increasing and continuous in v and satisfies the full-range condition. Throughout the following we focus on deterministic measures, which we may identify with the corresponding assignment $\boldsymbol{y} = (\boldsymbol{y}(1), \boldsymbol{y}(2))$.

Figure 2 illustrates the set of profiles $\boldsymbol{v} = (\boldsymbol{v}(1), \boldsymbol{v}(2))$ and, for each such profile, identifies the assignment(s) $\boldsymbol{y} = (\boldsymbol{y}(1), \boldsymbol{y}(2))$ implemented by that profile. The two lines, identifying profiles that make either $x = 1$ or $x = 2$ indifferent between the two elements of Y, form the boundaries of four closed (and hence overlapping on the boundaries) regions, whose union is the set $\boldsymbol{B}(Y)$ of profiles $\boldsymbol{v}$. All assignments $\boldsymbol{y} \in Y^X$ are implementable, but only the constant assignments $\boldsymbol{y} = (1, 1)$ and $\boldsymbol{y} = (2, 2)$ are strongly implementable.

[29] Throughout the following discussion we impose Assumption 3 and, therefore, suppose that the principal's utility is strictly increasing in the transfer received from the agent. As noted in Guesnerie and Laffont (1984), there is no reason to suppose that the participation constraint should be binding if this assumption fails.

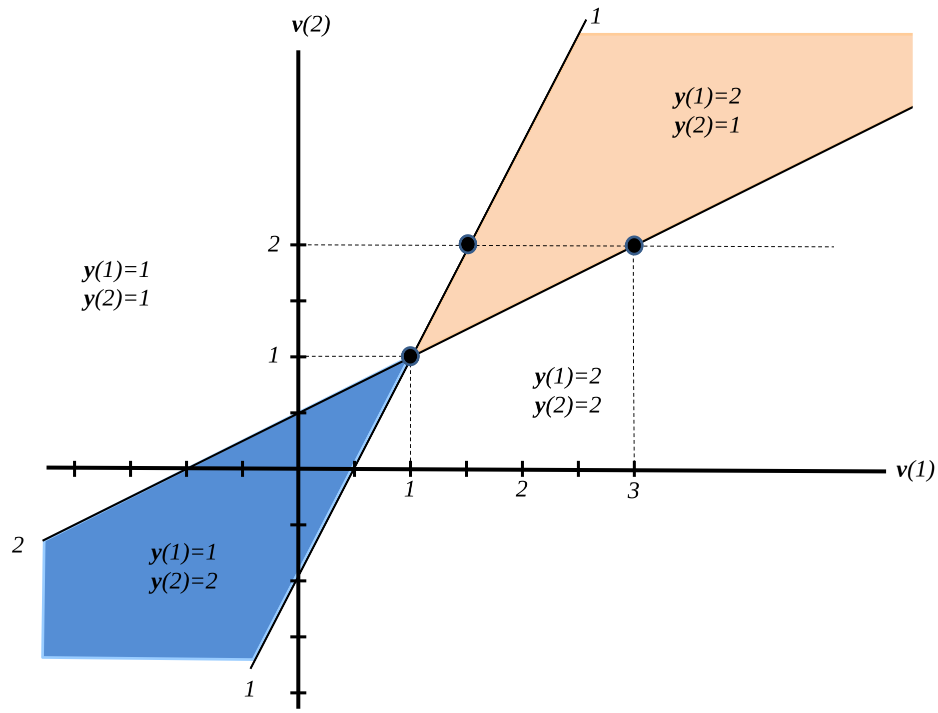

Figure 2: Illustration of the assignments $\boldsymbol{y}$ implemented by various profiles $\boldsymbol{v}$, the set $\boldsymbol{I}(Y)$ of implementable profiles (colored or shaded areas, including the boundary) and the feasible set for the principal's nonlinear pricing problem (the portion of the shaded areas for which $\boldsymbol{v}(2) \leq 2$) in Example 2. The profile $\hat{\boldsymbol{v}} = (1,1)$ is both the smallest profile implementing $\boldsymbol{y} = (2,1)$ and the largest profile implementing $\boldsymbol{y} = (1,2)$. As a consequence, neither of these two assignments is strongly implementable. The principal's optimum implements $\boldsymbol{y} = (1,2)$ while leaving both participation constraints slack.

The set of implementable tariffs $\boldsymbol{I}(Y)$ is the (blue and orange, or dark and light) shaded area in Figure 2, including the boundaries. This is immediate from Remark 6 upon observing that these tariffs are the ones implementing assignments that are onto Y.

All tariffs with $\boldsymbol{v}(2) \leq 2$ satisfy $\Phi \boldsymbol{v} \geq \underline{\boldsymbol{u}}$, whereas tariffs in the shaded area of Figure 2 with $\boldsymbol{v}(2) > 2$ lead to a violation of agent 1's participation constraint. Hence, the set $\{\boldsymbol{v} \in \boldsymbol{I}(Y) : \boldsymbol{v} \geq \Psi \underline{\boldsymbol{u}}\}$ appearing in the nonlinear pricing problem (38) is given by that portion of the shaded area in Figure 2 for which $\boldsymbol{v}(2) \leq 2$.

As the principal's utility function is strictly increasing in the payment v, there are only four candidates for a deterministic solution to the principal's problem: she could implement either $\boldsymbol{y} = (2,2)$ or $\boldsymbol{y} = (2,1)$ by choosing $\boldsymbol{v} = (3,2)$, she could implement $\boldsymbol{y} = (1,1)$ by choosing $(1.5,2)$, or she could implement $\boldsymbol{y} = (1,2)$ by choosing $\boldsymbol{v} = (1,1)$. Now, suppose the principal's utility function is

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

Then it is a straightforward calculation that among those four candidates, choosing $\boldsymbol{v} = (1, 1)$ to implement $\boldsymbol{y} = (1, 2)$ maximizes the principal's expected utility. The resulting utility profile for the agent is $\boldsymbol{u} = (1, 1)$, so that the participation constraint for neither type of agent binds in the unique solution to the principal's problem. ■

Example 2 features common values, in the sense that the principal cares directly about which type of the agent obtains which decision. This is an essential ingredient in the construction of the example: In the absence of such common values any change in tariff that changes the implemented assignment from $\boldsymbol{y} = (1, 2)$ to $\boldsymbol{y} = (2, 1)$ affects the principal's utility only through the change in tariff, ensuring that the principal would welcome the attendant increase in tariff from implementing $\boldsymbol{y} = (2, 1)$ with the tariff $\boldsymbol{v} = (3, 2)$ rather than implementing $\boldsymbol{y} = (1, 2)$ with the tariff $\boldsymbol{v} = (1, 1)$.

We say that the principal-agent model has private values if the principal's payoff function π does not depend on x, i.e., we can rewrite the principal's utility function as $\hat{\pi} : Y \times \mathbb{R} \to \mathbb{R}$. As we have just suggested, with private values no counterpart to Example 2 can be constructed:

Proposition 11. *Let Assumptions 1 and 3 hold and let the principal-agent model have private values. Then in any solution to the principal's problem, the participation constraint is binding for some type of agent.*

Appendix A.18 contains the proof. The key idea is that any $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ feasible in the principal's problem corresponds to a pairwise stable outcome satisfying the participation constraint $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ in a suitably constructed balanced matching model (cf. Remark 8). We can then apply the result in Corollary 7 to obtain a minimum (in the set of buyer profiles) pairwise stable outcome, in which the participation constraint binds for some type of buyer. The principal can implement this outcome, which features the same induced distribution ν over decisions as the one that we started from. The private-values assumption ensures that this leads to a strictly higher payoff for the principal than any feasible outcome in which the participation constraint is not binding.

5.4 Exclusion

Our formulation of the principal-agent model in Section 5.1 does not include an explicit outside option for the agent; rather it simply insists that the principal must respect the agent's participation constraint. It is clear, though, that in the presence of an outside option the principal may sometimes prefer to exclude some agent type(s) by designing a tariff that induces them to choose their outside option (Jullien, 2000). Here we show how to incorporate the possibility of exclusion into our model, explain why this leaves our existence result (Proposition 9) unchanged, and demonstrate that in the absence of quasilinearity or private values the principal might sometimes find it advantageous to "bribe" some type of the agent to be excluded.

To model the agent's outside option, we follow a strategy analogous to that used to incorporate non-participation in the matching model. Given a principal-agent model $(X, Y, \phi, \mu, \pi, \underline{u})$ satisfying Assumptions 1 and 3, we let $Y_0 = Y \cup \{y_0\}$, where the outside option y_0 is in the metric space containing Y, but is isolated from it, and extend the definition of the generating function ϕ to a function ϕ_0 on $X \times Y_0 \times \mathbb{R}$ satisfying Assumption 1 and

$$\phi_0(x, y_0, 0) = \underline{\boldsymbol{u}}(x). \tag{39}$$

Hence, in the absence of a transfer ($v = 0$), agent types choosing the outside option y_0 receive their reservation utility $\underline{\boldsymbol{u}}(x)$. Similarly, we extend the definition of the principal's utility function π to a function π_0 on $X \times Y_0 \times \mathbb{R}$ satisfying Assumption 3 and 5

$$\pi_0(x, y_0, v) = \underline{\pi}(v)$$

for some function $\underline{\pi} : \mathbb{R} \to \mathbb{R}$, with $\underline{\pi}(0)$ then specifying the principal's utility from not trading.

We will refer to $(X, Y_0, \phi_0, \mu, \pi_0, \underline{\boldsymbol{u}})$ as the principal-agent model with exclusion. Because we have supposed that Assumptions 1 and 3 carry over from $(X, Y, \phi, \mu, \pi, \underline{u})$ to $(X, Y_0, \phi_0, \mu, \pi, \underline{\boldsymbol{u}})$, it is immediate from Proposition 9 that the principal-agent model with exclusion has a solution $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ in which $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other. Further, because any such solution respects the participation constraint $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$, it satisfies the constraint that the principal cannot charge the agent for choosing the outside option.[30]

Corollary 8. *Let Assumptions 1 and 3 hold. The principal-agent model with exclusion has a solution* $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ *satisfying* $\boldsymbol{v}(y_0) \leq 0$.

Provided that the participation constraint binds for some type of agent in a solution to the principal-agent model with exclusion, we must have $\boldsymbol{v}(y_0) = 0$, and hence no agent is paid for nonparticipation. As the extension of the principal's payoff function to Y_0 preserves private values, this will be the case whenever the underlying principal-agent model satisfies the private value condition. Similarly, whenever the agent's utility function in the underlying principal-agent model is quasilinear and the specification of $\phi_0(x, y_0, v)$ is also quasilinear (i.e., we have $\phi_0(x, y_0, v) = \underline{\boldsymbol{u}}(x) - v$), then the principal-agent model with exclusion will satisfy quasilinearity, so that as in Jullien's quasilinear model of exclusion there is no loss of generality to restrict the principal to tariffs satisfying $\boldsymbol{v}(y_0) = 0$ (Jullien, 2000, footnote 7).[31]

If the participation constraint does not hold with equality for any type of the agent in a solution to the principal-agent model with exclusion, then such a solution might satisfy $\boldsymbol{v}(y_0) < 0$. There are two ways in which this might come about. The first possibility is that no type of the agent is excluded, but, as in Example 2, all types of the agent obtain strictly higher utility than their reservation utility. In this case, the optimal $(\boldsymbol{u}, \boldsymbol{y})$ can also be implemented by a (non-implementable) tariff $\boldsymbol{v}$ satisfying $\boldsymbol{v}(y_0) = 0$. The second, more interesting, case is that some excluded type receives the strictly positive payment $-\boldsymbol{v}(y_0)$ as a reward for not taking up any of the decisions in Y. The following example illustrates this can indeed occur.

Example 3. Let $X = \{1, 2\}$, let $Y = \{1\}$, and let $\mu(1) = \mu(2) = 1/2$. There are thus two equally likely types of agents, and the principal has the option of either assigning decision 1

[30] Using the obvious notation for the inverse generating function and implementation map in the model with exclusion, the formal argument is this: If $\boldsymbol{u}$ and $\boldsymbol{v}$ implement each other, the participation constraint implies $\boldsymbol{v} \leq \Psi_0 \underline{\boldsymbol{u}}$. Therefore, we have $\boldsymbol{v}(y_0) \leq \psi_0(y_0, x, \underline{\boldsymbol{u}}(x))$ for all $x \in X$. From (39), the later expression is equal to zero.

[31] Strong implementability of the optimal decision function in the principal-agent model (without exclusion) does not imply that the participation constraint holds as an equality in the principal-agent model with exclusion. Example 3 below (with only one decision in the absence of exclusion, so that strong implementability is immediate) provides an illustration.

to an agent (hereafter “interacting with the agent”) or excluding the agent by making him choose the outside option $y_0 = 0$.

The agents' utilities are given by

$$\phi_0(1,1,v) = 1 - v \qquad \phi_0(1,0,v) = -\frac{1}{2}v$$
$$\phi_0(2,1,v) = 2 - v \qquad \phi_0(2,0,v_0) = -2v,$$

and hence $\underline{\boldsymbol{u}}(1) = \underline{\boldsymbol{u}}(2) = 0$. The principal's utility is given by

$$\pi_0(1,1,v) = b + v \qquad \pi_0(1,0,v) = v$$
$$\pi_0(2,1,v) = v - c \qquad \pi_0(2,0,v_0) = v,$$

so that $\underline{\pi} = 0$. The parameter $b > 0$ is a benefit the principal obtains from interacting with an agent of type 1 and $c > 0$ is a corresponding cost of interacting with an agent of type 2. Now suppose that the principal's optimum involves interacting with agent 1 and excluding agent 2, as will be the case whenever both b and c are sufficiently large. Then the optimal tariff is $\boldsymbol{v}(1) = 2/3 = -\boldsymbol{v}(0)$. Hence, the principal pays agent 2 to stay out of the market. ■

6 Single Crossing

For unidimensional principal-agent models in which the agent's utility function is quasilinear, assuming the agent's willingness to pay to be strictly supermodular leads to a sharp characterization of (strongly) implementable assignments: an assignment is (strongly) implementable if and only if it is increasing (Rochet (1987), also see Vohra (2011, Theorem 4.2.5)). Similarly, for unidimensional matching models with perfectly transferable utility, assuming that the surplus function is strictly supermodular implies that all stable full matches feature positive assortative matching (Becker, 1973).

In this section we show that these results carry over to our setting with imperfectly transferable utility. The only change required is to replace the assumption of strict supermodularity with the assumption that the generating function satisfies a strict single crossing condition.

Assumption 4. *The sets X and Y are compact intervals in $\mathbb{R}$. The generating function ϕ satisfies the strict single crossing condition:*

$$\phi(x_1, y_2, v_2) \geq \phi(x_1, y_1, v_1) \Longrightarrow \phi(x_2, y_2, v_2) > \phi(x_2, y_1, v_1) \tag{40}$$

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

A quasilinear generating function $\phi(x, y, v) = f(x, y) - v$ satisfies the strict single crossing condition if and only if $f(x, y)$ is strictly supermodular.[32]

[32] Under quasilinearity, the strict single crossing condition (40) becomes

$$f(x_1, y_2) - f(x_1, y_1) \geq v_2 - v_1 \Longrightarrow f(x_2, y_2) - f(x_2, y_1) > v_2 - v_1.$$

This is obviously implied by the strict supermodularity condition $f(x_2, y_2) - f(x_2, y_1) > f(x_1, y_2) - f(x_1, y_1)$, while choosing $v_2 - v_1 = f(x_1, y_2) - f(x_1, y_1)$ ensures that strict single crossing implies supermodularity.

We begin by considering matching models satisfying Assumption 4 and then show how the results obtained for this case can be leveraged into a corresponding result for implementable assignments. The results we obtain generalize previous results for principal-agent models without quasilinear preferences by Guesnerie and Laffont (1984) and for matching models with imperfectly transferable utility by Legros and Newman (2007). The former impose a smoothness condition on the generating function and restrict attention to piecewise continuously differentiable assignments. The latter consider a model with a finite number of agents and show that their generalized increasing differences condition, which is equivalent to our strict single crossing condition, ensures that stable matches are positive assortative.

6.1 Positive Assortative Matching

We consider balanced matching models $(X, \mathcal{Y}, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfying Assumptions 1 and 4. Given that X and Y are compact intervals in the reals it will be convenient to identify the measures μ, ν, and λ with their distribution functions, denoted by F_μ, G_ν, and H_λ. Let λ^* be the unique match satisfying

$$H_{\lambda^*}(x, y) = \min\{F_\mu(x), G_\nu(y)\} \quad \text{for all } (x, y) \in X \times Y. \tag{41}$$

Following Galichon (2016, Chapter 4) we refer to λ^* as the positive assortative match.

When both F_μ and G_ν are continuous and strictly increasing, the positive assortative match is obtained by matching each agent with his or her uniquely determined counterpart on the other side of the market who has the same "rank" in the type distribution (as determined by the quantile functions F^{-1} and G^{-1}). Note that, in general, the positive assortative match need not be deterministic but will be so when μ is atomless (Galichon, 2016, Lemma 4.2). This provides us with the condition in the following proposition ensuring that the pairwise stable match is not only unique but also deterministic.

Proposition 12. *Let Assumptions 1 and 4 hold and the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be balanced. Then the positive assortative match λ^* is the unique pairwise stable match for all initial conditions (x_1, u_1). Further, if μ is absolutely continuous with respect to Lebesgue measure, then λ^* is deterministic.*

Proof. Proposition 6 ensures that there exists a pairwise stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ with $\boldsymbol{u}$ and $\boldsymbol{v}$ implementing each other and satisfying the initial condition (x_1, u_1).

Suppose the set $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ is comonotonic, that is, for (x_1, y_1) and $(x_2, y_2) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ we have $x_2 > x_1 \Rightarrow y_2 \geq y_1$ and $y_2 > y_1 \Rightarrow x_2 \geq x_1$. Proposition 5.1 then implies that supp (λ) is comonotonic. From Theorem 3 in Dhaene, Denuit, Goovaerts, Kaas, and Vyncke (2002), λ then satisfies (41) and therefore is the positive assortative match λ^*. If μ is absolutely continuous with respect to Lebesgue measure, then F_μ is continuous and λ^* is deterministic (Galichon, 2016, Lemma 4.2).

It remains to verify that the strict single crossing condition (40) in Assumption 4 implies that $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ is comonotonic. It suffices to show that there does not exist $(x_1, y_1), (x_2, y_2) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ satisfying $x_2 > x_1$ and $y_1 > y_2$. To show this, observe that (because $\boldsymbol{v}$ implements $\boldsymbol{u}$) from Lemma 2 we have that $(x_1, y_1), (x_2, y_2) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ implies

$$\phi(x_1, y_1, \boldsymbol{v}(y_1)) \geq \phi(x_1, y_2, \boldsymbol{v}(y_2))$$
$$\phi(x_2, y_2, \boldsymbol{v}(y_2)) \geq \phi(x_2, y_1, \boldsymbol{v}(y_1)).$$

With $x_2 > x_1$ and $y_1 > y_2$, the first of these inequalities and (40) imply $\phi(x_2, y_1, \boldsymbol{v}(y_1)) > \phi(x_2, y_2, \boldsymbol{v}(y_2))$, contradicting the second inequality. ■

Extending Proposition 12 to show that the unique pairwise stable match λ^* is also the unique stable match requires the existence of a pairwise stable outcome $(\lambda^*, \boldsymbol{u}, \boldsymbol{v})$ satisfying the participation constraints $\boldsymbol{u} \geq \underline{\boldsymbol{u}}$ and $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$. This isn't guaranteed. For an extreme counterexample, it may be that there is no pair of agents capable of generating individually rational payoffs (that is, $\underline{\boldsymbol{u}}(x) > \phi(x, y, \underline{\boldsymbol{v}}(y))$ holds for all (x, y)), obviously implying that in the unique stable outcome all agents are unmatched. Suppose, however, that for all $(x, y) \in X \times Y$, we have

$$\underline{\boldsymbol{u}}(x) < \phi(x, y, \underline{\boldsymbol{v}}(y)) \tag{42}$$

and consider a stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$. If there were unmatched types in Y (that is, $\text{supp}\ (\nu - \lambda_Y) \neq \emptyset$), then we could conclude from (26) that there exists $\hat{y} \in \text{supp}\ (\nu)$ such that $\boldsymbol{v}(\hat{y}) = \underline{\boldsymbol{v}}(\hat{y})$ holds. Using (29) and (42) this implies $\boldsymbol{u}(x) > \underline{\boldsymbol{u}}(x)$ for all $x \in \text{supp}\ (\mu)$, which in turn implies (from (25)) that there exist no unmatched types in X (that is, $\text{supp}\ (\mu - \lambda_X) = \emptyset$). As in a balanced match there are no matches featuring a strictly positive measure of unmatched agents on one side of the market but not on the other, we may thus conclude that λ is a full match. As every stable outcome featuring a full match is also pairwise stable, Proposition 12 then implies:

Corollary 9. *Let Assumptions 1 and 4 hold, let the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be balanced and let (42) hold. Then the positive assortative match λ^* is the unique stable match.*

Similar arguments, though with more tedious notation, show that if Assumptions 1 and 4 hold, then in any stable match, all *matched* agents are matched positive assortatively.

6.2 Increasing Assignments

It is a familiar result that implementable assignments must be increasing if a strict single crossing condition holds (e.g., Fudenberg and Tirole, 1991, Theorem 7.2, p. 260). Therefore, the main challenge in proving the following result is to show that every increasing assignment can be implemented with any initial condition. To obtain this, we build on Proposition 12 to show that for every increasing assignment the deterministic match associated with it can arise as the unique pairwise stable match in a suitably defined matching model.

Proposition 13. *Let Assumptions 1 and 4 hold. Then an assignment $\boldsymbol{y}$ is implementable if and only if it is increasing. In addition, every implementable assignment is strongly implementable.*

Proof. Suppose the assignment $\boldsymbol{y}$ is implementable. Then there exist $\boldsymbol{u}$ and $\boldsymbol{v}$ implementing each other, such that $(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ holds for all $x \in X$ (Proposition 4.1). Because $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ is comonotonic (cf. the proof of Proposition 12), this implies that $\boldsymbol{y}$ is increasing.

Fix an increasing assignment $\boldsymbol{y}$ and an initial condition (x_1, u_1). We construct a balanced matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$: Let μ be the restriction of Lebesgue measure on X to the Borel sets, and let ν be the pushforward of μ through $\boldsymbol{y}$ (which is well-defined because an increasing function $\boldsymbol{y}$ is measurable). The reservation utilities $\underline{\boldsymbol{u}}$ and $\underline{\boldsymbol{v}}$ will play no role, and so we can take $\underline{\boldsymbol{u}} \equiv 0 \equiv \underline{\boldsymbol{v}}$.

Let λ^* denote the positive assortative match for the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. From Proposition 12, λ^* is deterministic, and we have $\lambda^* = \lambda_{\boldsymbol{y}}$. Applying Proposition 12, we obtain that there exists $(\boldsymbol{u}, \boldsymbol{v})$ such that $(\lambda_{\boldsymbol{y}}, \boldsymbol{u}, \boldsymbol{v})$ is a pairwise stable outcome with $\boldsymbol{u}(x_1) = u_1$. From Proposition 6 we may take $\boldsymbol{u}$ and $\boldsymbol{v}$ to implement each other.

We complete the argument by showing that $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$. It suffices to show that for every $x \in X$, $(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ (Proposition 4). From Proposition 5.1, we have $\mathrm{supp}(\lambda_{\boldsymbol{y}}) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$. Fix a value $x \in X$. If $\boldsymbol{y}$ is continuous at x, then we immediately have $(x, \boldsymbol{y}(x)) \in \mathrm{supp}(\lambda_{\boldsymbol{y}})$ (since otherwise $\lambda_{\boldsymbol{y}}(\tilde{X} \times Y) = 0$ for some neighborhood $\tilde{X}$ of x, a contradiction). If $\boldsymbol{y}$ is not continuous at x, then the increasing function $\boldsymbol{y}$ must take an upward jump at x, and we have $(x, \boldsymbol{y}(x)) \in [\lim_{\tilde{x} \nearrow x} \boldsymbol{y}(x), \lim_{\tilde{x} \searrow x} \boldsymbol{y}(x)] \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$. The final inclusion follows from the facts that for each $y' \in [\lim_{\tilde{x} \nearrow x} \boldsymbol{y}(x), \lim_{\tilde{x} \searrow x} \boldsymbol{y}(x)]$ there exists $x' \in X$ such that $(x', y') \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ (because, from Lemma 2, $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ coincides with the graph of the argmax-correspondence $\boldsymbol{Y}_{\boldsymbol{v}}$, which is non-empty valued) and that $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ is comonotonic (cf. the proof of Proposition 12), which implies $x' = x$. ■

Recall from Section 2.4 that in the absence of quasilinearity an assignment may be implementable without being strongly implementable. Proposition 13 shows that strict crossing precludes this possibility. It follows that strict single crossing is a sufficient condition for the participation constraint to bind in any solution to the principal-agent model (Proposition 10).

Remark 9 (Single Crossing vs. Strict Single Crossing)**.** Say that the generating function satisfies the single crossing condition if the final inequality in (40) is weak. Under this weaker condition there may be (pairwise) stable matches that are different from the positive assortative match λ^* and non-increasing assignments may be implementable (as can be easily see by considering the trivial quasilinear example in which the generating function is given by $\phi(x, y, v) = -v$). However, under otherwise identical assumptions it remains true that in a balanced matching model the positive assortative match λ^* is pairwise stable for all initial conditions, that every balanced matching model satisfying condition (42) has a stable outcome featuring the match λ^*, and that every increasing assignment $\boldsymbol{y}$ is strongly implementable. Proving this is more tedious under single crossing than under strict single crossing as an extra step is required in the proof of Proposition 12 to show that the support of λ^* is contained in $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ for every pairwise stable outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ with $\boldsymbol{u}$ and $\boldsymbol{v}$ implementing each other. ■

7 Discussion

Strong implementability has appeared in our analysis in several places. This section first offers a condition that generalizes quasilinearity and suffices for strong implementability. We then touch briefly on two extensions of the analysis, and offer some concluding comments.

7.1 A Sufficient Condition for Strong Implementability

The absence of income effects under quasilinear preferences leads immediately to the translational invariance noted in Section 2.4, which in turn implies strong implementability. A weaker form of translational invariance suffices to obtain the conclusion of strong implementability of all implementable assignments. Specifically, consider a generating function satisfying Assumption 1. Fix any type $x_0 \in X$ and any profile $\boldsymbol{v}$. Then for any $t_0 \in \mathbb{R}$ we can find a uniquely determined profile $\hat{\boldsymbol{v}}$ such that $\phi(x_0, y, \boldsymbol{v}(y)) - \phi(x_0, y, \hat{\boldsymbol{v}}(y)) = t_0$ holds for all $y \in Y$. The optimal decisions of type x_0 when maximizing again the tariff $\boldsymbol{v}$ are then identical to the optimal decisions when maximizing against $\hat{\boldsymbol{v}}$. Further, the same will be true for any other type $x_1 \in X$ provided that there exists t_1 such that $\phi(x_1, y, \boldsymbol{v}(y)) - \phi(x_1, y, \hat{\boldsymbol{v}}(y)) = t_1$ holds for all y. Therefore, every implementable assignment is strongly implementable if the generating function satisfies for any x_0, x_1, y and y' and for any v, v', $\hat{v}$ and $\hat{v}'$,

$$
\begin{gathered}
[\phi(x_0, y, v) - \phi(x_0, y', v')] = [\phi(x_0, y, \hat{v}) - \phi(x_0, y', \hat{v}')] \\
\Rightarrow \\
[\phi(x_1, y, v) - \phi(x_1, y', v')] = [\phi(x_1, y, \hat{v}) - \phi(x_1, y', \hat{v}')].
\end{gathered}
\tag{43}
$$

Condition (43) ensures more than strong implementability. It also implies that (as in the quasilinear case) every match λ that is pairwise stable for *some* initial condition is pairwise stable for *all* initial conditions. This is so because, under condition (43), moving from a tariff $\boldsymbol{v}$ to any "translated tariff" $\hat{\boldsymbol{v}}$ leaves the argmax-correspondence unchanged: $\mathrm{supp}(\lambda) \in \Gamma_{\Phi \mathrm{v}, \mathrm{v}}$ implies $\mathrm{supp}(\lambda) \in \Gamma_{\Phi \hat{\mathrm{v}}, \hat{\mathrm{v}}}$.

We note that condition (43) embodies *no* restriction on the preferences of a single agent type x_0 over (y, v) pairs beyond the weak regularity properties implied by Assumption 1, and hence allows arbitrary income effects. Rather, condition (43) imposes a restriction *across* types, demanding that whatever change in tariff is required to preserve all utility differences for one type will also preserve all utility differences for any other type. Condition (43) holds, of course, if the characteristic function is quasilinear. More generally, it is satisfied if the characteristic function takes the form $\phi(x, y, v) = f(x, y) - h(y, v)$.

7.2 Stochastic Contracts in the Principal-Agent Model

In the principal-agent model with quasilinear utility it is well-known that the principal may benefit from offering stochastic rather then deterministic contracts to screen different agent types (cf. Strausz, 2006, for extensive discussion). In general, a stochastic contract corresponds to an incentive compatible direct mechanism which specifies, for every type of the agent, a lottery over transfers and decisions. To explain how stochastic contracts can be embedded in our model, it will be easier to begin with the case in which transfers are taken to be deterministic.

Fix a principal-agent model $(X, Y, \phi, \mu, \nu, \pi, \underline{\boldsymbol{u}})$ satisfying Assumptions 1 and 3 and let ΔY be the set of probability measures over the set Y, with typical element ζ. We equip the set ΔY with the topology of weak convergence, and note that ΔY is then a compact metric space (with the Prokhorov metric).

We can then extend the definitions of the payoff functions by taking the appropriate expectations:

$$
\begin{aligned}
\phi_\Delta(x, \zeta, v) &= \int_Y \phi(x, y, v) d\zeta(y) \\
\pi_\Delta(x, \zeta, v) &= \int_Y \pi(x, y, v) d\zeta(y),
\end{aligned}
$$

thereby obtaining a principal-agent model $(X, \Delta Y, \phi_\Delta, \mu, \pi_\Delta, \underline{\boldsymbol{u}})$ in which the set of possible decisions is given by ΔY rather than Y and a tariff assigns a transfer to every probability measure $\zeta \in \Delta Y$ rather than to every decision y. This extended principal-agent model satisfies Assumptions 1 and 3.[33] Consequently, our version of the taxation principle (Remark 1) as well as all the results from Section 5 continue to hold.

If both ϕ and π are quasilinear, then the restriction to deterministic *transfers* is without loss of generality, as both the agent's and the principal's preferences only depend on the expected transfer. In the general case this is not so, raising the question whether we can incorporate stochastic transfers in our model. That we can do so is not immediately obvious because the duality theory developed in Sections 2 and 3 of this paper hinges on a tariff being a map into the real numbers. However, while doing so would be redundant for deterministic contracts, there is nothing in the formal structure of the model which prevents us from supposing that decisions y include the specification of a monetary transfer.[34] Therefore, the

[33]We have already noted that ΔY is a compact metric space. It is obvious that ϕ_Δ and π_Δ inherit the requisite monotonicity properties and the full range condition from ϕ and π. Continuity is less obvious. From the definition of weak convergence and the fact that for fixed x and v, the function $\phi(x, y, v) : Y \to \mathbb{R}$ is continuous on a compact set, we can conclude that if the sequence $(\zeta_n)_{n=1}^\infty$ converges (weakly) to the limit ζ, then

$$
\int_Y \phi(x, y, v) d\zeta_n(y) \to \int_Y \phi(x, y, v) d\zeta(y). \tag{44}
$$

This in turn implies that ϕ_Δ is continuous: Suppose we have a sequence $(x_n, \zeta_n, v_n)_{n=1}^\infty$ converging to (x, ζ, v) (pointwise in the first and third arguments, and in the sense of weak convergence in the second). Notice that the set $\{v_n\}_{n=1}^\infty$ is contained in a compact subset $\tilde{\mathbb{R}}$ of $\mathbb{R}$. Then for any ε, there exists a sufficiently large N such that, for all $n \geq N$,

$$
\begin{aligned}
& \left| \int_Y \phi(x_n, y, v_n) d\zeta_n(y) - \int_Y \phi(x, y, v) d\zeta(y) \right| \\
\leq\ & \left| \int_Y \phi(x_n, y, v_n) d\zeta_n(y) - \int_Y \phi(x, y, v) d\zeta_n(y) \right| + \left| \int_Y \phi(x, y, v) d\zeta_n(y) - \int_Y \phi(x, y, v) d\zeta(y) \right| \\
\leq\ & \left| \int_Y (\phi(x_n, y, v_n) - \phi(x, y, v) d\zeta_n(y)) \right| + \frac{\varepsilon}{2} \\
\leq\ & \int_Y \frac{\varepsilon}{2} d\zeta_n(y) + \frac{\varepsilon}{2} \\
\leq\ & \varepsilon,
\end{aligned}
$$

where the first appearance of $\varepsilon/2$ follows from (44) and the second follows from the uniform continuity of the function ϕ on the compact set $X \times Y \times \tilde{\mathbb{R}}$. A similar argument applies to establish continuity of π_Δ.

[34]For example, let $q \in [0, \bar{q}]$ be the quantity of some good. Ordinarily, we would take $Y = [0, \bar{q}]$ and then

same construction that we have described above—replacing the set Y by the set ΔY—allows us to introduce stochastic transfers into the model with the only salient restriction being that any randomization over payments that comes on top of the deterministic transfer v is restricted to a compact set of probability distributions.

7.3 Moral Hazard in the Principal-Agent Model

We have considered adverse-selection principal-agent models. Following Myerson (1982), Laffont and Tirole (1993), Laffont and Martimort (2002, Section 7.1), Kadan, Reny, and Swinkels (2017) and others, one might extend the model to encompass moral hazard. The recipe for incorporating moral hazard is similar to that for stochastic contracts. We offer a simple illustration.

Suppose the agent must choose an effort level $e \in [0, 1]$ that induces a probability mass function $f(z, e)$ with support on the finite set Z, from which an output z is realized. The principal cannot observe the agent's effort. Once again, we can view the agent as choosing a decision y and paying a transfer $\boldsymbol{v}(y)$ to the principal. A decision y now is a function $\boldsymbol{w} : Z \to [\underline{w}, \overline{w}]$ identifying, for each output level z, the wage $\boldsymbol{w}(z) \in [\underline{w}, \overline{w}]$ paid by the principal to the agent if output z is realized. The agent's utility from wage w, output z, effort level e and transfer v is given by $u(x, e, w - v)$, while the principal's utility is $z - (w - v)$.

The set X is again a compact set of agent types. We take the set Y to be the set of functions $\boldsymbol{w} : Z \to [\underline{w}, \overline{w}]$. Then we let

$$\phi(x, \boldsymbol{w}, v) = \max_{e \in [0,1]} \sum_{z \in Z} u(x, e, \boldsymbol{w}(z) - v) f(z, e).$$

We let $\mathcal{E}(x, \boldsymbol{w})$ be the set of maximizers of this problem, and let the principal's utility be

$$\pi(x, \boldsymbol{w}, v) = \max_{e \in \mathcal{E}(x, \boldsymbol{w})} \sum_{z \in Z} (z - (\boldsymbol{w}(z) - v)) f(z, e).$$

Assuming that u and f are continuous, it follows from Berge's maximum theorem that ϕ is continuous, and hence Assumption 1 is satisfied. The function $\pi(x, \boldsymbol{w}, v)$ is upper semicontinuous. We would again have Assumptions 1 and 3 satisfied, except that the function π is only semicontinuous. However, this suffices for an argument analogous to that of Section 5.

One might want to generalize this illustration in many ways, including allowing an infinite set of possible outputs and relaxing the bounds on the function $\boldsymbol{w}$. Our results will apply as long as attention is restricted to circumstances in which the set Y can reasonably be taken to be compact.

---

suppose that a monopolistic seller (the principal) with utility function $\pi(x, q, v)$ designs a tariff specifying payments $\boldsymbol{v}(q)$ for all possible quantities that a consumer (the agent) with preferences described by the utility function $\phi(x, q, v)$ might want to buy. Instead, we may take $\hat{Y} = [0, \bar{q}] \times [0, \bar{t}]$ and suppose that the seller prices bundles $(q, t) \in Y$, consisting of a quantity q of the good and a rebate $t \in [0, \bar{t}]$ that the consumer receives if he buys the bundle (q, t) at price $\boldsymbol{v}(q, t)$. Setting $\hat{\phi}(x, y, v) = \phi(x, q, v - t)$ and $\hat{\pi}(x, y, v) = \pi(x, q, v - t)$ for $y = (q, t)$ then yields a principal-agent model $(X, \hat{Y}, \hat{\phi}, \mu, \hat{\pi}, \underline{\boldsymbol{u}})$ that satisfies Assumption 1 and 3 if the original model $(X, Y, \phi, \mu, \pi, \underline{\boldsymbol{u}})$ does so and describes the same underlying economic environment.

7.4 Conclusion

We have introduced and studied a duality relationship that provides a characterization of implementable profiles and assignments suitable for 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.

Throughout our analysis we have eschewed smoothness assumptions, as these play no role for the duality structure and are not required for the existence and characterization results pursued here. However, much of the power of conjugate duality stems from the inherent smoothness properties of convex functions, and many of the more useful implications of generalized conjugate duality for the quasilinear case—ranging from the familiar integral representation of implementable utility profile (e.g. Myerson, 1979) to results asserting the uniqueness and determinateness of stable matchings (e.g. Chiappori, McCann, and Nesheim, 2010)—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, McCann and Zhang (2017) use the implementation duality to show how the conditions from Figalli, Kim, and McCann (2011), under which the principal's problem can be reduced to a convex maximization program, can be extended to the non-quasilinear case. We also believe that (under suitable differentiability assumptions) it will be possible to extend the type-assignment approach developed in Nöldeke and Samuelson (2007), which rests on the inverse relationship between a pair of profiles implementing each other and their associated argmax-correspondences, to characterize optimal bunches in principal-agent models satisfying the strict single crossing condition.

A number of extensions suggest themselves. First, much is known about the structure of the set of stable outcomes in matching models with a finite number of agents (Roth and Sotomayor, 1990, Chapter 9), including connectedness and comparative static properties, that one might want to extend to the current setting. Second, as suggested by our discussion of stochastic contracts and moral-hazard in the principal-agent model, our compactness assumption on Y is sometimes restrictive because it is natural to allow for unbounded Y. Similarly, the assumption that the type space X is compact is violated in some applications in finance (such as Glosten, 1989) in which normally distributed types are considered.[35] Third, the implementation relationships studied here also appear in economic contexts different from the ones we have considered, with possible applications ranging from the characterization of hedonic pricing equilibria (cf. Chiappori, McCann, and Nesheim, 2010, in the quasilinear case) to the development of new econometric techniques for discrete-choice random-utility models (Bonnet, Galichon, and Shum, 2017). Finally, while Galois connections have played little role in economic theory so far, their appearance in the study of information aggregation (under the guise of a residual mapping) in Chambers and Miller (2011) and in the study of preference aggregation (Monjardet, 1978, 2007), suggest that further applications may by found in areas far removed from the implementation duality.

[35] In the quasilinear case Bardsley (2017) provides an illuminating duality-based analysis of principal-agent models that avoids compactness assumptions.

Appendix

A.1 Implementability and Direct Mechanisms

Let $\mathbb{R}^X$ be the set of functions from X to $\mathbb{R}$. Then $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ (note that here $\boldsymbol{u}$ is not required to be bounded) is implementable by an incentive compatible direct mechanism if there exists $\boldsymbol{t} \in \mathbb{R}^X$ such that the feasibility conditions $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x))$ and the incentive compatibility conditions $\phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x)) \geq \phi(x, \boldsymbol{y}(\hat{x}), \boldsymbol{t}(\hat{x}))$ hold for all $x, \hat{x} \in X$. Similarly, letting $\mathbb{R}^Y$ be the set of functions from Y to $\mathbb{R}$, we may define $(\boldsymbol{v}, \boldsymbol{x}) \in \mathbb{R}^Y \times X^Y$ to be implementable by an incentive compatible direct mechanism if there exists $\boldsymbol{t} \in \mathbb{R}^Y$ such that $\boldsymbol{v}(y) = \psi(y, \boldsymbol{x}(y), \boldsymbol{t}(y))$ and $\psi(y, \boldsymbol{x}(y), \boldsymbol{t}(y)) \geq \psi(y, \boldsymbol{x}(\hat{y}), \boldsymbol{t}(\hat{y}))$ hold for all $y, \hat{y} \in Y$.

Lemma 7. *Let Assumption 1 hold.*

[7.1] $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ is implementable by an incentive compatible direct mechanism if and only if $\boldsymbol{u} \in \boldsymbol{B}(X)$ and there exists $\boldsymbol{v} \in \boldsymbol{B}(Y)$ implementing $(\boldsymbol{u}, \boldsymbol{y})$.

[7.2] $(\boldsymbol{v}, \boldsymbol{x}) \in \mathbb{R}^Y \times X^Y$ is implementable by an incentive compatible direct mechanism if and only if $\boldsymbol{v} \in \boldsymbol{B}(Y)$ and there exists $\boldsymbol{u} \in \boldsymbol{B}(X)$ implementing $(\boldsymbol{v}, \boldsymbol{x})$.

Proof of Lemma 7. We prove Lemma 7.1; the proof of Lemma 7.2 is analogous.

It is immediate from the revelation principle that if $(\boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{B}(X) \times Y^X$ is implemented by $\boldsymbol{v} \in \boldsymbol{B}(Y)$ then $(\boldsymbol{u}, \boldsymbol{y})$ is implementable by an incentive compatible direct mechanism. Indeed, upon setting $\boldsymbol{t}(x) = \boldsymbol{v}(\boldsymbol{y}(x))$ for all $x \in X$, conditions (3) and (4) imply $\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x)) \geq \phi(x, \boldsymbol{y}(\hat{x}), \boldsymbol{t}(\hat{x}))$ for all $x, \hat{x} \in X$.

Conversely, suppose that $(\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^X \times Y^X$ is implementable by an incentive compatible direct mechanism, so that there exists $\boldsymbol{t} \in \mathbb{R}^X$ such that

$$\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x)) \geq \phi(x, \boldsymbol{y}(\hat{x}), \boldsymbol{t}(\hat{x})) \tag{45}$$

$$\boldsymbol{t}(x) = \psi(\boldsymbol{y}(x), x, \boldsymbol{u}(x)) \geq \psi(\boldsymbol{y}(x), \hat{x}, \boldsymbol{u}(\hat{x})) \tag{46}$$

hold for all $x, \hat{x} \in X$, where the equality in (46) follows from (45) because ϕ and ψ are inverse and the inequality in (46) follows from (45) upon reversing the roles of x and $\hat{x}$ in the inequality $\boldsymbol{u}(x) \geq \phi(x, \boldsymbol{y}(\hat{x}), \boldsymbol{t}(\hat{x}))$ and using, again, that ϕ and ψ are inverse.

First, we establish that $\boldsymbol{u}$ is bounded. Fix $\hat{x} \in X$. The inequality in (45) ensure that for all $x \in X$,

$$\boldsymbol{u}(x) \geq \phi(x, \boldsymbol{y}(\hat{x}), \boldsymbol{t}(\hat{x})) \geq \min_{\tilde{x} \in X} \phi(\tilde{x}, \boldsymbol{y}(\hat{x}), \boldsymbol{t}(\hat{x})) =: \underline{u} \in \mathbb{R},$$

where the minimum $\underline{u}$ exists because X is compact and ϕ continuous. Next, using (46) we have,

$$\boldsymbol{t}(x) = \psi(\boldsymbol{y}(x), x, \boldsymbol{u}(x)) \geq \psi(\boldsymbol{y}(x), \hat{x}, \boldsymbol{u}(\hat{x})) \geq \min_{y \in Y} \psi(y, \hat{x}, \boldsymbol{u}(\hat{x})) =: \underline{t} \in \mathbb{R},$$

for all $x \in X$, where the minimum $\underline{t}$ exists because Y is compact and ψ continuous. Using the equality in (45) and that ϕ is strictly decreasing in its third argument, we then have, for all $x \in X$,

$$\boldsymbol{u}(x) = \phi(x, \boldsymbol{y}(x), \boldsymbol{t}(x)) \leq \phi(x, \boldsymbol{y}(x), \underline{t}) \leq \max_{\tilde{x} \in X, \tilde{y} \in Y} \phi(\tilde{x}, \tilde{y}, \underline{t}) =: \overline{u} \in \mathbb{R},$$

where the maximum $\overline{u}$ exists because X and Y are compact and ϕ continuous. We thus have $\underline{u} \leq \boldsymbol{u}(x) \leq \overline{u}$ for all $x \in X$, which implies $\boldsymbol{u} \in \boldsymbol{B}(X)$. From the equality in (46), $\boldsymbol{t}$ is bounded, too.

Second, we show there exists $\boldsymbol{v} \in \boldsymbol{B}(Y)$ implementing $(\boldsymbol{u}, \boldsymbol{y})$. We can fix a value $\overline{v} \in \mathbb{R}$ such that $\phi(x, y, \overline{v}) \leq \underline{u}$ holds for all $(x, y) \in X \times Y$. Now let

$$\boldsymbol{v}(y) = \begin{cases} \boldsymbol{t}(x) & \text{if } y = \boldsymbol{y}(x) \text{ for some x} \in \text{X} \\ \overline{v} & \text{otherwise.} \end{cases}$$

If there exist $x, \hat{x} \in X$ and $y \in Y$ with $y = \boldsymbol{y}(x) = \boldsymbol{y}(\hat{x})$, then the incentive constraints in (45) imply $\boldsymbol{t}(x) = \boldsymbol{t}(\hat{x})$. Therefore $\boldsymbol{v}(y)$ is well-defined for all $y \in Y$ and, because $\boldsymbol{t}$ is bounded, we have $\boldsymbol{v} \in \boldsymbol{B}(Y)$. Finally, using (45), it is immediate from the construction of $\boldsymbol{v}$ that we have

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

for all $(x, y) \in X \times Y$, so that $\boldsymbol{v}$ implements $(\boldsymbol{u}, \boldsymbol{y})$. ■

A.2 Proof of Lemma 1

First, 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 $\varepsilon > 0$. We have to establish that there exists $\delta > 0$ such that

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

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 $\delta \in (0, 1)$ and $x \in X$, we then have

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

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

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

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

$$\mid \psi(y, x, \tilde{\boldsymbol{u}}(x)) - \psi(y, x, \boldsymbol{u}(x)) \mid < \varepsilon' \text{ for all } x \in X \text{ and } y \in Y \Longrightarrow$$
$$\sup_{y \in Y} \left| \sup_{x \in X} \psi(y, x, \tilde{\boldsymbol{u}}(x)) - \sup_{x \in X} \psi(y, x, \boldsymbol{u}(x)) \right| \leq \varepsilon' < \varepsilon,$$

which gives $\|\Psi\tilde{\boldsymbol{u}} - \Psi\boldsymbol{u}\| < \varepsilon$, as desired.

Second, let $\mathcal{V} \subset \boldsymbol{B}(Y)$ be 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. The argument for Ψ is analogous. ■

A.3 Proof of Proposition 2

It is immediate from the definitions that $\boldsymbol{I}(X) \subseteq \Phi\boldsymbol{B}(Y)$. Hence, to establish the first statement in (13) we need to show that the image $\Phi\boldsymbol{v}$ of any profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is implementable and continuous. The remaining statement in (13) follows by an analogous argument.

Given any profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$, let $s_{\boldsymbol{v}} = \sup_{y \in Y} \boldsymbol{v}(y)$ denote its supremum and $i_{\boldsymbol{v}} = \inf_{y \in Y} \boldsymbol{v}(y)$ its infimum. These are finite because $\boldsymbol{v}$ is bounded. Let $E_{\boldsymbol{v}} = \{(y, v) \in Y \times \mathbb{R} \mid v \geq \boldsymbol{v}(y)\}$ denote the epigraph of $\boldsymbol{v}$, and let $Z_{\boldsymbol{v}} = \{(y, v) \in Y \times \mathbb{R} \mid s_{\boldsymbol{v}} \geq v \geq \boldsymbol{v}(y)\}$. Observe that the set $Z_{\boldsymbol{v}} \subset E_{\boldsymbol{v}}$ is bounded, contains the graph of $\boldsymbol{v}$ and is contained in $[i_{\boldsymbol{v}}, s_{\boldsymbol{v}}] \times Y$, which is a compact set (because Y is compact).

Lemma 8. *Let Assumption 1 hold. If a profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is lower semicontinuous, then it implements $\Phi\boldsymbol{v}$, $\Phi\boldsymbol{v}$ is continuous, and the argmax correspondence $\boldsymbol{Y}_{\boldsymbol{v}}$ is nonempty- and compact-valued and upper hemicontinuous. Analogously, if a profile $\boldsymbol{u} \in \boldsymbol{B}(X)$ is lower semicontinuous, then it implements $\Psi\boldsymbol{u}$, $\Psi\boldsymbol{u}$ is continuous, and the argmax correspondence $\boldsymbol{X}_{\boldsymbol{u}}$ is nonempty- and compact-valued and upper hemicontinuous.*

Proof of Lemma 8. If $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is lower semicontinuous, then its epigraph $E_{\boldsymbol{v}}$ is closed and so is $Z_{\boldsymbol{v}}$. As $Z_{\boldsymbol{v}}$ is contained in the compact set $[i_{\boldsymbol{v}}, s_{\boldsymbol{v}}] \times Y$ it follows that $Z_{\boldsymbol{v}}$ is compact. As the generating function ϕ is continuous, a solution to the problem

$$\max_{(y,v) \in Z_{\boldsymbol{v}}} \phi(x, y, v) \tag{47}$$

thus exists for all $x \in X$ by Weierstrass' extreme value theorem. As ϕ is continuous and $Z_{\boldsymbol{v}}$ is compact, it follows from Berge's maximum theorem (Ok, 2007, p. 306) that the profile $\boldsymbol{u} \in \boldsymbol{B}(X)$ defined by $\boldsymbol{u}(x) = \max_{(y,v) \in Z_{\boldsymbol{v}}} \phi(x, y, v)$ for all $x \in X$ is continuous and the correspondence mapping X into $\operatorname{argmax}_{(y,v) \in Z_{\boldsymbol{v}}} \phi(x, y, v)$ is compact valued and upper hemicontinuous.

We next show that $\boldsymbol{u} = \Phi\boldsymbol{v}$ and that $\boldsymbol{v}$ implements $\boldsymbol{u}$. As the graph of $\boldsymbol{v}$ is contained in $Z_{\boldsymbol{v}}$, we have

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

On the other hand, because ϕ is strictly decreasing in its third argument any solution to (47) lies on the graph of $\boldsymbol{v}$, implying that for every $x \in X$, there exists $\boldsymbol{y}(x) \in Y$ such that

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

holds. This ensures that the suprema in the definition of $\Phi\boldsymbol{v}$ are attained and that $\boldsymbol{v}$ implements $\Phi\boldsymbol{v} = \boldsymbol{u}$. Finally, the compact-valuedness and upper hemicontinuity of $\boldsymbol{Y}_{\boldsymbol{v}}$ are implied by the same properties of the correspondence mapping X into $\operatorname{argmax}_{(y,v) \in Z_{\boldsymbol{v}}} \phi(x, y, v)$. ■

Continuation of the Proof of Proposition 2. It remains to consider the case in which $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is not lower semicontinuous. Let $\bar{\boldsymbol{v}}$ be the lower semicontinuous hull of $\boldsymbol{v}$, i.e., the greatest element of the family of lower semicontinuous functions from Y to $\mathbb{R}$ majorized by $\boldsymbol{v}$.[36] (The existence of $\bar{\boldsymbol{v}}$ is assured, cf. Penot (2013, Proposition 1.21).) As $\boldsymbol{v}$ is bounded,

[36]The lower semicontinuous hull of a function is also known as its lsc regularization or its lower closure (Rockafellar and Wets, 1998, p. 14).

so is $\bar{\boldsymbol{v}}$, i.e., we have $\bar{\boldsymbol{v}} \in \boldsymbol{B}(Y)$. From Lemma 8 the profile $\bar{\boldsymbol{v}}$ implements $\Phi\bar{\boldsymbol{v}}$, which is continuous. It remains to show that $\Phi\bar{\boldsymbol{v}} = \Phi\boldsymbol{v}$ holds. Because the epigraph $E_{\bar{\boldsymbol{v}}}$ of $\bar{\boldsymbol{v}}$ is the closure of the epigraph $E_{\boldsymbol{v}}$ of $\boldsymbol{v}$ (Penot, 2013, Proposition 1.21), we have that $Z_{\bar{\boldsymbol{v}}}$ is the closure of $Z_{\boldsymbol{v}}$. Therefore,

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

and thus (because ϕ is decreasing in its third argument) we have $\sup_{y\in Y} \phi(x,y,\boldsymbol{v}(y)) = \max_{y\in Y} \phi(x,y,\bar{\boldsymbol{v}}(y))$ for all $x \in X$, which is the desired result. ■

A.4 Proof of Corollary 3

We prove Corollary 3.1; 3.2 is analogous. First, if $\boldsymbol{\Psi}u$ implements $(\boldsymbol{u},\boldsymbol{y})$, then $(\boldsymbol{u},\boldsymbol{y})$ is implementable by definition.

Second, let $(\boldsymbol{u},\boldsymbol{y}) \in \boldsymbol{B}(X)\times Y^X$ be implementable. Then $\boldsymbol{u}$ implements $\Psi\boldsymbol{u}$ (Corollary 2). It remains to show that $\Psi\boldsymbol{u}$ implements $(\boldsymbol{u},\boldsymbol{y})$. As $\Psi\boldsymbol{u}$ is implementable, it implements $\Phi\Psi\boldsymbol{u}$ (Corollary 2). Proposition 3.1 gives $\boldsymbol{u} = \Phi\Psi\boldsymbol{u}$, so that we have shown that $\Psi\boldsymbol{u}$ implements $\boldsymbol{u}$. To show that $\Psi\boldsymbol{u}$ also implements $\boldsymbol{y}$, we proceed as follows:

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 which we obtain $\Psi\boldsymbol{u} = \Psi\Phi\tilde{\boldsymbol{v}}$. From the first inequality in (10) in Corollary 1.1, we have $\tilde{\boldsymbol{v}} \geq \Psi\Phi\tilde{\boldsymbol{v}}$ and thus $\tilde{\boldsymbol{v}} \geq \Psi\boldsymbol{u}$. Now suppose that $\Psi\boldsymbol{u}$ does not implement $\boldsymbol{y}$. Because $\Psi\boldsymbol{u}$ implements $\boldsymbol{u}$ there then exists $(\hat{x},\hat{y}) \in X \times Y$ such that

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

where the last inequality uses $\tilde{\boldsymbol{v}} \geq \Psi\boldsymbol{u}$ 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. ■

A.5 Proof of (18)–(19) in Remark 6

We prove (19); (18) is analogous. First, suppose the profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ is implementable. Then $\boldsymbol{v}$ implements and is implemented by $\boldsymbol{u} = \Phi\boldsymbol{v}$ (Corollary 3), implying that both $X_{\boldsymbol{u}}$ and $Y_{\boldsymbol{v}}$ are nonempty valued. Further, from Lemma 2 the correspondences are inverses of each other, and hence must be onto.

Second, suppose that $Y_{\boldsymbol{v}}$ is nonempty valued and onto. Then $\boldsymbol{v}$ implements the profile $\boldsymbol{u} = \Phi\boldsymbol{v}$ (because $Y_{\boldsymbol{v}}$ is nonempty valued) and for all $\hat{y} \in Y$ there exists $\hat{x} \in X$ such that $\boldsymbol{u}(\hat{x}) = \phi(\hat{x},\hat{y},\boldsymbol{v}(\hat{y}))$ holds (because $Y_{\boldsymbol{v}}$ is onto), which is equivalent to $\boldsymbol{v}(\hat{y}) = \psi(\hat{y},\hat{x},\boldsymbol{u}(\hat{x}))$. As $\boldsymbol{v}$ implements $\boldsymbol{u}$ we have $\boldsymbol{u}(x) \geq \phi(x,\hat{y},\boldsymbol{v}(\hat{y}))$ for all $x \in X$, which is equivalent to $\boldsymbol{v}(\hat{y}) \geq \psi(\hat{y},x,\boldsymbol{u}(x))$ for all $x \in X$. Combining the equality and the inequality for $\boldsymbol{v}(\hat{y})$ we have $\boldsymbol{v}(\hat{y}) = \max_{x\in X} \phi(\hat{y},x,\boldsymbol{u}(x))$. As this holds for all $\hat{y} \in Y$, it follows that $\boldsymbol{u}$ implements $\boldsymbol{v}$, so that $\boldsymbol{v}$ is implementable. ■

A.6 Proof of Corollary 4.

We prove statements [4.1]–[4.3], with the proofs of the corresponding statements for $\boldsymbol{I}(Y)$ being analogous.

[4.1] Consider a sequence $(\boldsymbol{u}_n)_{n=1}^{\infty}$ of profiles in $\boldsymbol{I}(X)$ converging to some $\boldsymbol{u}^* \in \boldsymbol{B}(X)$. We want to show that $\boldsymbol{u}^*$ is implementable. For all $n \in \mathbb{N}$, let $\boldsymbol{v}_n = \Psi \boldsymbol{u}_n$. Because Ψ is continuous (Lemma 1), the sequence $(\boldsymbol{v}_n)_{n=1}^{\infty}$ converges to $\boldsymbol{v}^* = \Psi \boldsymbol{u}^*$. Corollary 3.1 implies that $\boldsymbol{v}_n$ implements $\boldsymbol{u}_n$, so that we have $\boldsymbol{u}_n = \Phi \boldsymbol{v}_n$ for all $n \in N$. Taking limits on both sides of this equation and using the continuity of Φ (Lemma 1), we obtain $\boldsymbol{u}^* = \Phi \boldsymbol{v}^*$. From Proposition 2 this establishes the implementability of $\boldsymbol{u}^*$, and hence that $\boldsymbol{I}(X)$ is closed. Next, suppose that the sequence $(\boldsymbol{u}_n)_{n=1}^{\infty}$ is in $\mathcal{U}_{\boldsymbol{y}} \subset \boldsymbol{I}(X)$. With the same construction of the sequence $(\boldsymbol{v}_n)_{n=1}^{\infty}$ as above, Corollary 3.1 then implies that $\boldsymbol{v}_n$ implements $\boldsymbol{y}$ for all n, so that

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

holds for all $x \in X$, $y \in Y$ and $n \in \mathbb{N}$. As the (uniform) convergence of $(\boldsymbol{v}_n)_{n=1}^{\infty}$ to $\boldsymbol{v}^*$ implies its pointwise convergence to the same limit and ϕ is continuous, the above inequalities imply

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

for all $x \in X$ and $y \in Y$. Therefore, $\boldsymbol{v}^*$ implements $\boldsymbol{y}$. As $\boldsymbol{v}^*$ also implements $\boldsymbol{u}^*$, this establishes $\boldsymbol{u}^* \in \mathcal{U}_{\boldsymbol{y}}$.

[4.2] Let $\mathcal{U} \subset \boldsymbol{I}(X)$ be bounded. Fix $\varepsilon > 0$. To show equicontinuity of $\mathcal{U}$, we establish that there exists $\delta > 0$ such that

$$\|\hat{x} - x\| < \delta \Longrightarrow \|\boldsymbol{u}(\hat{x}) - \boldsymbol{u}(x)\| < \varepsilon \tag{48}$$

for all $\hat{x}, x \in X$ and $\boldsymbol{u} \in \mathcal{U}$.

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

$$\|\hat{x} - x\| < \delta \Longrightarrow \|\phi(\hat{x}, y, v) - \phi(x, y, v)\| < \varepsilon \tag{49}$$

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

Consider any $\boldsymbol{u} \in \mathcal{U}$. From Corollary 3, the profile $\boldsymbol{v} = \Psi \boldsymbol{u} \in \mathcal{V}$ implements $\boldsymbol{u}$. Let $\tilde{y} \in \boldsymbol{Y}_{\boldsymbol{v}}(x)$ and $\hat{y} \in \boldsymbol{Y}_{\boldsymbol{v}}(\hat{x})$. We then have

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

implying

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

where the outer inequalities are from (49) and the fact that $\underline{v} \leq \boldsymbol{v}(y) \leq \bar{v}$ holds for all $y \in Y$. Consequently, we have $\|\boldsymbol{u}(\hat{x}) - \boldsymbol{u}(x)\| < \varepsilon$, thus establishing (48).

[4.3] This follows from Corollary 4.2 and an application of the Arzela-Ascoli theorem (Ok, 2007, p. 264). ■

A.7 Proof of Lemma 4

We prove Lemma 4.1; the proof for Lemma 4.2 is analogous.

Let $\mathfrak{U} \subset \boldsymbol{I}(X)$ be a closed sublattice of $\boldsymbol{B}(X)$ for which

$$U_x = \{\boldsymbol{u} \in \mathfrak{U} | \boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)\}.$$

is non-empty for all $x \in X$. For the current proof, the important observation is that if $\boldsymbol{y} \in Y^X$ is strongly implementable, then one such set is $\mathcal{U}_{\boldsymbol{y}}$, which is a subset of $\boldsymbol{I}(X)$ (by definition), closed (Corollary 4.1), and, by Lemma 3, a sublattice of $\boldsymbol{B}(X)$, with the strong implementability of $\boldsymbol{y}$ ensuring that $\{\boldsymbol{u} \in \mathcal{U}_{\boldsymbol{y}} | \boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)\}$ is nonempty for all $x \in X$.

Let

$$S = \{\boldsymbol{u} \in \mathfrak{U} \mid \boldsymbol{u} \geq \underline{\boldsymbol{u}}\}.$$

We proceed in two steps. The first step establishes that there exists $\hat{\boldsymbol{u}} \in S$ satisfying $\hat{\boldsymbol{u}}(x) = \underline{\boldsymbol{u}}(x)$ for some $x \in X$. The second step then completes the argument by showing that S has a minimal element.

Step 1: Pick an arbitrary $x_0 \in X$ and $\boldsymbol{u}_0$ in U_{x_0} (recalling that U_x is non-empty for all $x \in X$). We construct a sequence $(x_n)_{n=1}^{\infty}$ in X and an associated sequence $(\boldsymbol{u}_n)_{n=1}^{\infty}$ of profiles in $\mathfrak{U}$, satisfying $\boldsymbol{u}_n \in U_{x_n}$ for all n, by the following recursion: Given $(x_{n-1}, \boldsymbol{u}_{n-1})$ with $\boldsymbol{u}_{n-1} \in U_{x_{n-1}}$, let $x_n \in \arg\min_{x \in X} [\boldsymbol{u}_{n-1}(x) - \underline{\boldsymbol{u}}(x)]$. Because both $\boldsymbol{u}_{n-1}$ (as an implementable profile) and $\underline{\boldsymbol{u}}$ (by assumption) are continuous and X is compact, such an x_n exists. Pick any $\hat{\boldsymbol{u}}_n \in U_{x_n}$. Define $\boldsymbol{u}_n = \boldsymbol{u}_{n-1} \vee \hat{\boldsymbol{u}}_n$. Because $\mathfrak{U}$ is a sublattice, we then have $\boldsymbol{u}_n \in \mathfrak{U}$. Because $\boldsymbol{u}_{n-1} \in U_{x_{n-1}}$ implies $\min_{x \in X} [\boldsymbol{u}_{n-1}(x) - \underline{\boldsymbol{u}}(x)] \leq 0$, we further have $\boldsymbol{u}_n(x_n) = \underline{\boldsymbol{u}}(x_n)$, implying $\boldsymbol{u}_n \in U_{x_n}$.

The sequence $(\boldsymbol{u}_n)_{n=1}^{\infty}$ is increasing by construction. It is also bounded above.[37] Therefore, it is bounded and thus equicontinuous (Corollary 4.2). Hence, $(\boldsymbol{u}_n)_{n=1}^{\infty}$, which is a sequence in the closed set $\mathfrak{U}$, has a limit point $\hat{\boldsymbol{u}} \in \mathfrak{U}$. Note that $\hat{\boldsymbol{u}}$ is continuous because it is implementable (Proposition 2).

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 the continuous function $\hat{\boldsymbol{u}}$ we have

$$\lim_{k \to \infty} \boldsymbol{u}_{n_k}(x_{n_k}) = \hat{\boldsymbol{u}}(x^*) \tag{50}$$

$$\lim_{k \to \infty} \boldsymbol{u}_{n_{k-1}}(x_{n_k}) = \hat{\boldsymbol{u}}(x^*). \tag{51}$$

As $\boldsymbol{u}_n(x_n) = \underline{\boldsymbol{u}}(x_n)$ holds for all n and $\underline{\boldsymbol{u}}$ is continuous, (50) implies

$$\hat{\boldsymbol{u}}(x^*) = \underline{\boldsymbol{u}}(x^*). \tag{52}$$

[37] By Assumption 1 and the continuity of the profile $\underline{\boldsymbol{u}}$, the profile $\underset{\sim}{\boldsymbol{v}} \in \boldsymbol{B}(Y)$ given by $\underset{\sim}{\boldsymbol{v}}(y) = \min_{x \in X} \psi(y, x, \underline{\boldsymbol{u}}(x))$ for all $y \in Y$ is well-defined. For any profile $\boldsymbol{v} \in \boldsymbol{B}(Y)$ satisfying $\boldsymbol{v}(\hat{y}) < \underset{\sim}{\boldsymbol{v}}(\hat{y})$ for some $\hat{y} \in Y$, we have $\phi(x, \hat{y}, \boldsymbol{v}(\hat{y})) > \underline{\boldsymbol{u}}(x)$ for all $x \in X$ by construction. For such $\boldsymbol{v}$, $\boldsymbol{u} = \Phi \boldsymbol{v}$ thus satisfies $\boldsymbol{u}(x) > \underline{\boldsymbol{u}}(x)$ for all $x \in X$, implying that $\boldsymbol{u}$ is not in $\cup_{x \in X} U_x$. By the order reversal property of the implementation map Φ it follows that $\overline{\boldsymbol{u}} = \Phi \underset{\sim}{\boldsymbol{v}}$ is an upper bound for $\cup_{x \in X} U_x$ and therefore an upper bound for $(\boldsymbol{u}_n)_{n=1}^{\infty}$.

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

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

for all $x \in X$, where we have used the continuity of $\underline{\boldsymbol{u}}$ and (51) to obtain the right side of the inequality. Taking account of (52) this implies

$$\hat{\boldsymbol{u}}(x) \geq \underline{\boldsymbol{u}}(x) \tag{53}$$

for all $x \in X$. Combining (52) and (53), we have established the desired result.

Step 2: As S contains $\hat{\boldsymbol{u}}$ satisfying $\hat{\boldsymbol{u}}(x) = \underline{\boldsymbol{u}}(x)$ for some $x \in X$, it is immediate that a minimum element $\boldsymbol{u}^*$ of S must satisfy $\boldsymbol{u}^*(x) = \underline{\boldsymbol{u}}(x)$ for the same $x \in X$. It remains to show that such a minimum element exists.

Given any $\bar{\boldsymbol{u}} \in S$, let

$$S_{\bar{\boldsymbol{u}}} = \{\boldsymbol{u} \in \mathfrak{U}\} \mid \bar{\boldsymbol{u}} \geq \boldsymbol{u} \geq \underline{\boldsymbol{u}}\}.$$

The set $S_{\bar{\boldsymbol{u}}}$ contains $\bar{\boldsymbol{u}}$ and hence is nonempty. Further, it is clearly bounded. As the intersection of two closed sets, the set $S_{\bar{\boldsymbol{u}}}$ is closed and as an intersection of two sublattices of $\boldsymbol{B}(X)$, it is a sublattice. With the set $S_{\bar{\boldsymbol{u}}}$ being a closed and bounded subset of $\boldsymbol{I}(X)$, it is compact (Corollary 4.3) and thus a complete sublattice of $\boldsymbol{B}(X)$.[38] The complete sublattice $S_{\bar{\boldsymbol{u}}}$ has a minimum element $\boldsymbol{u}^*$, which clearly is also the minimum element of S. ■

A.8 Stable Outcomes in Finite-Support Matching Models

We complete the argument from Remark 7. Let $\mathcal{X} = \{x \in X \mid x = x_i \text{ for some } i \in I\}$ and $\mathcal{Y} = \{y \in Y \mid y = y_j \text{ for some } j \in J\}$ denote the supports of the type distributions in the finite-support matching model. For $x \in \mathcal{X}$ let $I(x) = \{i \in I \mid x_i = x\}$ and for $y \in \mathcal{Y}$ let $J(y) = \{j \in J \mid y_j = y\}$ Consider now a stable outcome $(\rho, u_1, \ldots, u_m, v_1, \ldots, v_n)$ for the matching model with a finite number of agents. Let $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ be arbitrary profiles in $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$. Given that equal treatment holds, setting

$$\boldsymbol{u}(x) = \begin{cases} u_i & \text{if } x \in I(x) \\ \tilde{\boldsymbol{u}} & \text{otherwise} \end{cases}$$

and

$$\boldsymbol{v}(y) = \begin{cases} v_j & \text{if } y \in J(y) \\ \tilde{\boldsymbol{v}} & \text{otherwise} \end{cases}$$

[38] The set $S_{\bar{\boldsymbol{u}}}$ is compact in the norm topology. A lattice is complete if and only if it is compact in the interval topology (Birkhoff, 1995, p. 250, Theorem 20). Compactness in the norm topology implies compactness in the interval topology, as any set open under the latter is also open under the former.

gives two well-defined profiles $\boldsymbol{u} \in \boldsymbol{B}(X)$ and $\boldsymbol{v} \in \boldsymbol{B}(Y)$. The measure λ has support contained in $\mathcal{X} \times \mathcal{Y}$ and on this set is given by

$$\lambda(x, y) = \sum_{i \in I(x)} \sum_{j \in J(y)} \rho(i, j).$$

With these definitions, it is straightforward to verify that $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a stable outcome for the finite-support matching model.

A.9 Proof of Proposition 5.3

Let $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ be a pairwise stable outcome for the balanced matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. Let $\mathcal{X} \subseteq X$ and $\mathcal{Y} \subseteq Y$ be the supports of μ and ν. Noticing that $\mathrm{supp}(\lambda) \subseteq \mathcal{X} \times \mathcal{Y}$ holds, every pair of profiles $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ that satisfy $\tilde{\boldsymbol{u}} = \boldsymbol{u}$ on $\mathcal{X}$ and $\tilde{\boldsymbol{v}} = \boldsymbol{v}$ on $\mathcal{Y}$ satisfy (24) and (29), implying that for any such pair $(\lambda, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}})$ is a pairwise stable outcome. It thus suffices to construct a pair of profiles satisfying $\tilde{\boldsymbol{u}} = \boldsymbol{u}$ on $\mathcal{X}$ and $\tilde{\boldsymbol{v}} = \boldsymbol{v}$ on $\mathcal{Y}$ that implement each other.

Because λ is a full match, for every $x \in \mathcal{X}$ there exists $y \in \mathcal{Y}$ with $(x, y) \in \mathrm{supp}(\lambda)$. (Otherwise we would have $\lambda_X(\tilde{X}) = 0$ for some neighborhood $\tilde{X}$ of x, a contradiction.) By (24) and (29) this implies that the restriction of the profile $\boldsymbol{v}$ to $\mathcal{Y}$ implements the restriction of the profile $\boldsymbol{u}$ to $\mathcal{X}$, that is,

$$\boldsymbol{u}(x) = \max_{y \in \mathcal{Y}} \phi(x, y, \boldsymbol{v}(y)), \quad \forall x \in \mathcal{X}.$$

Similarly, for every $y \in \mathcal{Y}$ there must exist $x \in \mathcal{X}$ with $(x, y) \in \mathrm{supp}(\lambda)$, so that (24) and (29) imply that restriction of $\boldsymbol{u}$ to $\mathcal{X}$ implements the restriction of $\boldsymbol{v}$ to $\mathcal{Y}$:

$$\boldsymbol{v}(y) = \max_{x \in \mathcal{X}} \psi(y, x, \boldsymbol{u}(x)), \quad \forall y \in \mathcal{Y}.$$

Now define the profile $\tilde{\boldsymbol{u}} \in \boldsymbol{B}(X)$ by

$$\tilde{\boldsymbol{u}}(x) = \max_{y \in \mathcal{Y}} \phi(x, y, \boldsymbol{v}(y)).$$

This profile satisfies $\tilde{\boldsymbol{u}} = \boldsymbol{u}$ on $\mathcal{X}$ (because the restriction of $\boldsymbol{v}$ to $\mathcal{Y}$ implements the restriction of $\boldsymbol{u}$ to $\mathcal{X}$). Further, it is implementable. Indeed, because $\boldsymbol{v}$ is bounded for sufficiently large $\breve{v}$ any profile $\hat{\boldsymbol{v}} \in \boldsymbol{B}(Y)$ of the form

$$\hat{\boldsymbol{v}}(y) = \begin{cases} \boldsymbol{v}(y) & \text{if } y \in \mathcal{Y} \\ \breve{v} & \text{otherwise} \end{cases}$$

implements $\tilde{\boldsymbol{u}}$. Now, let $\tilde{\boldsymbol{v}} = \Psi \tilde{\boldsymbol{u}}$. As $\tilde{\boldsymbol{u}}$ is implementable, we then have that $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ implement each other (Proposition 3.1). It remains to show that $\tilde{\boldsymbol{v}} = \boldsymbol{v}$ holds on $\mathcal{Y}$. But this follows upon noting that $\tilde{\boldsymbol{u}} = \boldsymbol{u}$ on $\mathcal{X}$ implies $\tilde{\boldsymbol{v}} \geq \boldsymbol{v}$ on $\mathcal{Y}$ on the one hand (because the restriction of $\boldsymbol{u}$ to $\mathcal{X}$ implements the restriction of $\boldsymbol{v}$ to $\mathcal{Y}$) and on the other hand we have $\tilde{\boldsymbol{v}} = \Psi\Phi\hat{\boldsymbol{v}}$, which implies (from Corollary 1.1) $\hat{\boldsymbol{v}} \geq \tilde{\boldsymbol{v}}$ and therefore, because $\hat{\boldsymbol{v}} = \boldsymbol{v}$ on $\mathcal{Y}$, the inequality $\boldsymbol{v} \geq \tilde{\boldsymbol{v}}$ on $\mathcal{Y}$. ■

A.10 Proof of Lemma 5

Suppose $\boldsymbol{y}$ is implementable and satisfies $\lambda = \lambda_{\boldsymbol{y}}$. From Proposition 4.1, the implementability of $\boldsymbol{y}$ implies that there exists $\boldsymbol{u}$ and $\boldsymbol{v}$ implementing each other such that the graph of $\boldsymbol{y}$ is contained in $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$. As the argmax correspondence $Y_{\boldsymbol{v}}$ is upper hemicontinuous (Corollary 2), its graph is closed. Hence, $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$, which coincides with the graph of $Y_{\boldsymbol{v}}$ (Lemma 2), also contains the closure of the graph of $\boldsymbol{y}$. Moreover, the closure of the graph of $\boldsymbol{y}$ contains the support of $\lambda_{\boldsymbol{y}}$ (otherwise, there is a point (x, y) with a neighborhood that does not intersect the graph of $\boldsymbol{y}$ and which receives positive measure under $\lambda_{\boldsymbol{y}}$, a contradiction to the definition of $\lambda_{\boldsymbol{y}}$ in (32)). We thus have $\text{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$, implying that λ is pairwise stable (Proposition 5.1 and 5.2).

Conversely, suppose the deterministic match λ is pairwise stable. From Proposition 5.3 the pairwise stability of λ implies that there exist $(\boldsymbol{u}, \boldsymbol{v})$ implementing each other such that $\text{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$. By Proposition 4.1 it remains to show that there exists a measurable assignment $\boldsymbol{y}$ with graph contained in $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$ satisfying $\lambda_{\boldsymbol{y}} = \lambda$. By definition of a deterministic match, there exists a measurable assignment $\boldsymbol{y}'$ such that $\lambda = \lambda_{\boldsymbol{y}'}$ holds. If the graph of $\boldsymbol{y}'$ is contained in the support of λ, then we are done upon setting $\boldsymbol{y} = \boldsymbol{y}'$. It remains to consider the case that the graph of $\boldsymbol{y}'$ is not contained in the support of λ.

We construct the assignment $\boldsymbol{y}$. Let $\mathcal{X}$ denote the support of μ. First, we note that $\lambda_{\boldsymbol{y}'}$ does not depend on the specification of $\boldsymbol{y}'$ outside the support of μ. In addition, it is straightforward to define the assignment $\boldsymbol{y}$ on $X \setminus \mathcal{X}$ so that $(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$ holds for all $x \in X \setminus \mathcal{X}$.[39] Now let $\tilde{X} = \{x \in \mathcal{X} \mid (x, \boldsymbol{y}'(x)) \notin \text{supp}(\lambda)\}$. The set $\tilde{X}$ is negligible (that is, contained in a subset of $\mathcal{X}$ with measure zero) by definition of $\lambda_{\boldsymbol{y}'}$. Hence, we can complete the specification of $\boldsymbol{y}$ by taking $\boldsymbol{y}$ to equal a measurable selection from $\boldsymbol{Y}_{\boldsymbol{v}}$ (cf. footnote 39) (and hence $(x, \boldsymbol{y}(x)) \in \Gamma_{\boldsymbol{u},\boldsymbol{v}}$) on a subset of $\mathcal{X}$ that contains $\tilde{X}$ and has measure zero, and taking $\boldsymbol{y}$ to equal $\boldsymbol{y}'$ (and hence $(x, \boldsymbol{y}(x)) \in \text{supp}(\lambda) \subseteq \Gamma_{\boldsymbol{u},\boldsymbol{v}}$) on the remainder of $\tilde{X}$. This construction ensures that the graph of $\boldsymbol{y}$ is contained in $\Gamma_{\boldsymbol{u},\boldsymbol{v}}$. It follows immediately from the definitions of $\lambda_{\boldsymbol{y}}$ and $\lambda_{\boldsymbol{y}'}$ that we further have $\lambda_{\boldsymbol{y}} = \lambda_{\boldsymbol{y}'}$. As $\lambda_{\boldsymbol{y}'} = \lambda$ holds by assumption, this implies $\lambda_{\boldsymbol{y}} = \lambda$, finishing the proof. ∎

A.11 Proof of Proposition 6

Let $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ be a balanced matching problem satisfying Assumption 1. Since this matching model is balanced, nothing is lost (and some convenience is gained) by taking μ and ν to be probability measures, which we hereafter maintain.

Let $(x_1, \ldots, x_n) \in X^n$ and $(y_1, \ldots, y_n) \in Y^n$ satisfy $y_1 = y_0$, where $y_0 \in Y$ is the agent appearing as part of the initial condition (y_0, v_0) in the statement of the Proposition. Define a measure μ_n on X by

$$\mu_n(\tilde{X}) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{x_k}(\tilde{X}), \tag{54}$$

[39] The properties of the correspondence $\boldsymbol{Y}_{\boldsymbol{v}}$ noted in Corollary 2 imply that this correspondence has a closed graph, ensuring that it is weakly measurable (Aliprantis and Border, 2006, Theorem 18.20 and Lemma 18.2), and hence has a measurable selection (Aliprantis and Border, 2006, Theorem 18.13) $\tilde{\boldsymbol{y}}$. Take $\boldsymbol{y}$ to equal $\tilde{\boldsymbol{y}}$ on $X \setminus \mathcal{X}$.

for measurable $\tilde{X} \subseteq X$ and define the measure ν_n on Y similarly by

$$\nu_n(\tilde{Y}) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{y_k}(\tilde{Y}) \tag{55}$$

for all measurable $\tilde{Y} \subseteq Y$.

Lemma 9. *Let Assumption 1 hold. The matching model $(X, Y, \phi, \mu_n, \nu_n, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ has a pairwise stable outcome $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ with profiles $\boldsymbol{u}_n$ and $\boldsymbol{v}_n$ that implement each other and that satisfy $\boldsymbol{v}_n(y_0) = v_0$.*

Proof of Lemma 9. We first construct an auxiliary balanced finite-support matching model $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underset{\sim}{\boldsymbol{u}}, \underset{\sim}{\boldsymbol{v}})$ satisfying Assumption 1 by (i) multiplying the measures μ_n and ν_n by n (so as to convert them into counting measures) and (ii) replacing the reservation utility profiles $\underline{\boldsymbol{u}}$ and $\underline{\boldsymbol{v}}$ by reservation utility profiles

$$\underset{\sim}{\boldsymbol{u}}(x) = \underline{u}, \quad \forall x \in X$$

and

$$\underset{\sim}{\boldsymbol{v}}(y) = \begin{cases} v_0 & \text{if } y = y_0 \\ \underline{u} & \text{otherwise,} \end{cases}$$

where $\underline{u}$ is sufficiently small as to ensure $\phi(x, y, \underline{u}) > \phi(x, y_0, v_0) > \underline{u}$ for all $x \in X$ and $y \in Y$.

Consider the matching model with a finite number of agents associated with $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underset{\sim}{\boldsymbol{u}}, \underset{\sim}{\boldsymbol{v}})$ (cf. Remark 7). By construction of $\underset{\sim}{\boldsymbol{u}}$ and $\underset{\sim}{\boldsymbol{v}}$, the inequalities $\phi(x_i, y_j, \underline{u}) > \phi(x_i, y_0, v_0) > \underline{u}$ hold for all $i, j \in \{1, \dots, n\}$. Because there are an equal number of buyers and sellers, these inequalities ensure that there are no unmatched agents in a stable outcome and similarly preclude the possibility that any seller with $y_k \neq y_0$ obtains her reservation utility in a stable outcome. Hence, it follows from Lemma 3 in Demange and Gale (1985) that this matching model with a finite number of agents has a stable outcome in which all buyers and sellers are matched and sellers with $y_k = y_0$ obtain their reservation utility. This implies (cf. Appendix A.8) that the finite-support matching model $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underset{\sim}{\boldsymbol{u}}, \underset{\sim}{\boldsymbol{v}})$ has a fully matched stable outcome $(\hat{\lambda}_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ satisfying the initial condition $\boldsymbol{v}(y_0) = v_0$. As any fully matched stable outcome is also pairwise stable and the pairwise stability conditions do not depend on the reservation utility profiles, the outcome $(\hat{\lambda}_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ is also pairwise stable for the finite-support matching model $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. Letting $\lambda_n = \hat{\lambda}_n / n$, it is obvious that $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ is a pairwise stable outcome for the matching model $(X, Y, \phi, \mu_n, \nu_n, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. Finally, from Proposition 5.3 we may assume that $\boldsymbol{u}_n$ and $\boldsymbol{v}_n$ implement each other, giving a pairwise stable outcome $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ satisfying all the conditions from the statement of the lemma. ■

Let $(x_n)_{n1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ be sequences in X and Y with $y_1 = y_0$ and such that the probability measures μ_n and ν_n defined in (54)–(55) converge weakly to μ, respectively ν.

The existence of such sequences is assured: for example, if all but x_1 and y_1 are obtained by taking sequences of independent random draws from the probability measures μ and ν, then with probability one we obtain sequences of measures μ_n and ν_n that converge weakly (as $n \to \infty$) to the measures μ and ν (e. g. , Villani, 2009, p. 64). For each n, the matching model $(X, Y, \phi, \mu_n, \nu_n, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ has a pairwise stable outcome $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ satisfying the properties mentioned in the statement of Lemma 9. Let $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)_{n=1}^{\infty}$ be a sequence of such outcomes. The following lemma establishes that this sequence has a limit point, which is the pairwise stable outcome we seek.

Lemma 10. *Let Assumption 1 hold. The sequence $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)_{n=1}^{\infty}$ has a subsequence converging (weakly in the case of the measures λ_n, and in norm for the profiles) to a pairwise stable outcome $(\lambda^*, \boldsymbol{u}^*, \boldsymbol{v}^*)$ of the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ that satisfies $\boldsymbol{v}^*(y_0) = v_0$.*

Proof of Lemma 10 Because each of the probability measures λ_n is defined on the compact (and hence separable) metric space $X \times Y$, the collection $\{\lambda_n\}_{n=1}^{\infty}$ is tight, and Prokhorov's theorem (Shiryaev, 1996, p. 318) ensures that there is a subsequence of $(\lambda_n)_{n=1}^{\infty}$ converging weakly to a probability measure λ^* on $X \times Y$. Further, as each λ_n is a full match, so is λ^*, that is, conditions (30)–(31) are preserved in the limit (see Villani, 2009, p.64). For convenience of notation, we assume that the sequence $(\lambda_n)_{n=1}^{\infty}$ itself converges to λ^*.

We show below that the sequences $(\boldsymbol{u}_n)_{n=1}^{\infty}$ and $(\boldsymbol{v}_n)_{n=1}^{\infty}$ are bounded. Because $\{\boldsymbol{u}_n\}_{n=1}^{\infty}$ and $\{\boldsymbol{v}_n\}_{n=1}^{\infty}$ are sets of implementable profiles, Corollary 4.2 then ensures that both of these sets are equicontinuous so that it follows from the Arzela-Ascoli theorem (Kelley, 1955, p. 233) that $(\boldsymbol{u}_n, \boldsymbol{v}_n)_{n=1}^{\infty}$ has a subsequence (which, for notational convenience, we take to be the sequence itself) converging to some limit $(\boldsymbol{u}^*, \boldsymbol{v}^*)$. As the sets of implementable profiles are closed (Corollary 4.1) it follows that $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$ are implementable. Further, the arguments in the proof of Corollary 4.1 show that $(\boldsymbol{u}^*, \boldsymbol{v}^*)$ implement each other. As $\boldsymbol{v}_n(y_0) = v_0$ holds for all n, we obtain $v^*(y_0) = v_0$. In light of Proposition 5 the desired result then follows provided that $\mathrm{supp}(\lambda^*) \subseteq \Gamma_{\boldsymbol{u}^*, \boldsymbol{v}^*}$ holds, that is, we need to establish

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

for all $(x, y) \in \mathrm{supp}(\lambda^*)$. The weak convergence of the sequence $(\lambda_n)_{n=1}^{\infty}$ to λ^* ensures that for every (x, y) in the support of λ^*, there is a sequence $(x_n, y_n)_{n=1}^{\infty}$, with each (x_n, y_n) in the support of λ_n, converging to (x, y). For each n and each $(x_n, y_n) \in \mathrm{supp}(\lambda_n)$, we have

$$\boldsymbol{u}_n(x_n) = \phi(x_n, y_n, \boldsymbol{v}_n(y_n)).$$

The convergence of the equicontinuous sequences $(\boldsymbol{u}_n)_{n=1}^{\infty}$ and $(\boldsymbol{v}_n)_{n=1}^{\infty}$ of continuous profiles to the continuous profiles $(\boldsymbol{u}^*, \boldsymbol{v}^*)$ then gives the result.

It remains to establish boundedness of the sequences $(\boldsymbol{u}_n)_{n=1}^{\infty}$ and $(\boldsymbol{v}_n)_{n=1}^{\infty}$. To do so, we first recall that in the pairwise stable outcome $(\lambda_n, \boldsymbol{u}_n, \boldsymbol{v}_n)$ of the nth matching model, the profiles $\boldsymbol{u}_n$ and $\boldsymbol{v}_n$ implement each other and (because $y_1 = y_0$) satisfy $\boldsymbol{v}_n(y_1) = v_0$. Hence, for each n and x, we have

$$\boldsymbol{u}_n(x) \geq \phi(x, y_1, v_0) \geq \min_{x \in X} \phi(x, y_1, v_0),$$

providing us with a lower bound for $(\boldsymbol{u}_n)_{n=1}^{\infty}$. Similarly, we note that some buyer is matched with seller y_1. The ability of any seller to match with this buyer puts a lower bound on $\boldsymbol{v}_n$. We cannot be sure which buyer is involved in such a match, but we know that the buyer in question receives utility $\phi(x, y_1, v_0)$, and so we have

$$\boldsymbol{v}_n(y) \geq \min_{x \in X} \psi(y, x, \phi(x, y_1, v_0)),$$

providing us with a lower bound for $(\boldsymbol{v}_n)_{n=1}^{\infty}$. By the order reversal property of the implementation maps (Corollary 1.2) the lower bound on $(\boldsymbol{u}_n)_{n=1}^{\infty}$ provides us with an upper bound on $(\boldsymbol{v}_n)_{n=1}^{\infty}$ and the lower bound on $(\boldsymbol{v}_n)_{n=1}^{\infty}$ provides us with an upper bound on $(\boldsymbol{u}_n)_{n=1}^{\infty}$. This ensures that the sequences $(\boldsymbol{u}_n)_{n=1}^{\infty}$ and $(\boldsymbol{v}_n)_{n=1}^{\infty}$ are bounded, finishing the proof. ■

This completes the proof of Proposition 6. ■

A.12 Proof of Corollary 5

Fix a matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfying Assumption 1. We construct an augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ as follows.

First, we augment the type spaces X and Y by adding dummy types x_0 and y_0, where x_0 and y_0 are elements of the metric spaces containing X and Y but are not contained in X or Y. We let $X_0 = X \cup \{x_0\}$ and $Y_0 = Y \cup \{y_0\}$.

Second, the reservation utility profiles $\underline{\boldsymbol{u}}_0$ and $\underline{\boldsymbol{v}}_0$ duplicate $\underline{\boldsymbol{u}}$ on X and $\underline{\boldsymbol{v}}$ on Y, with $\underline{\boldsymbol{u}}(x_0) = \underline{\boldsymbol{v}}(y_0) = 0$.

Third, we let the generating function ϕ_0 equal ϕ on $X \times Y \times \mathbb{R}$, and then extend ϕ_0 to $X_0 \times Y_0 \times \mathbb{R}$ by defining

$$\begin{aligned} \phi_0(x, y_0, v) &= \underline{\boldsymbol{u}}(x) - v \\ \phi_0(x_0, y, v) &= \underline{\boldsymbol{v}}(y) - v \\ \phi_0(x_0, y_0, v) &= \quad\;\; - v. \end{aligned}$$

We let ψ_0 denote the inverse generating function associated with ϕ_0. Note that ψ_0 satisfies $\psi_0(y, x_0, u) = \underline{\boldsymbol{v}}(y) - u$, indicating that any type of seller y receives her reservation utility $\underline{\boldsymbol{v}}(y)$ when matching with a buyer x_0 who receives her reservation utility $\underline{\boldsymbol{u}}_0(x_0) = 0$, thus mirroring the utility obtained by a buyer of any type x who matches with y_0.

Fourth, the measure μ_0 duplicates μ on the set X, and attaches mass $\nu(Y) + 1$ to the isolated point x_0. Similarly, the measure ν_0 duplicates ν on the set Y, and attaches mass $\mu(X) + 1$ to the isolated point y_0. Note that $\mu_0(X_0) = \nu_0(Y_0) = 1 + \mu(X) + \nu(Y)$ holds, and so the matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ is balanced.

The augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ features continuous reservation utility profiles and satisfies Assumption 1: the sets X_0 and Y_0 are compact because X and Y are so, and the generating function ϕ_0 satisfies the full range condition and is continuous because the profiles $\underline{\boldsymbol{u}}$ and $\underline{\boldsymbol{v}}$ used in the construction of the extension of ϕ are (by assumption) continuous.

With any full match λ_0 for $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ we associate the match λ for $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ obtained by restricting λ_0 to $X \times Y$. Vice versa, we can extend any match

λ for $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ to a full match λ_0 for $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ by assigning the masses of unmatched agents to the dummy agents and matching the remaining masses of the dummy agents with each other. That is, we associate with λ the uniquely defined measure λ_0 satisfying

$$\begin{aligned}\lambda_0(\tilde{X} \times \{y_0\}) &= \mu(\tilde{X}) - \lambda_X(\tilde{X}) \\ \lambda_0(\{x_0\} \times \tilde{Y}) &= \nu(\tilde{Y}) - \lambda_Y(\tilde{Y})\end{aligned}$$

for all measurable $\tilde{X} \subseteq X$ and $\tilde{Y} \subseteq Y$, and

$$\lambda_0(\{x_0\} \times \{y_0\}) = 1 + \lambda(X \times Y).$$

We say that a full outcome $(\lambda, \boldsymbol{u}_0, \boldsymbol{v}_0)$ for $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ and an outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ for $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ are associated if (i) λ_0 and λ are associated, (ii) $\boldsymbol{u}$ is the restriction of $\boldsymbol{u}_0$ to X, and (iii) $\boldsymbol{v}$ is the restriction of $\boldsymbol{v}_0$ to Y.

Because the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ is balanced, we can invoke Proposition 6 to conclude that it has a pairwise stable outcome $(\lambda_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$ satisfying $\boldsymbol{u}_0(x_0) = 0$. The proof is then completed by the "if" direction of the following lemma. (The "only-if" direction of the lemma will be required in the proof of the subsequent Proposition 8.)

Lemma 11. *Let the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfy Assumption 1. Then $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a stable outcome of $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ if and only if it is associated with a pairwise stable outcome $(\lambda_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$, satisfying $\boldsymbol{u}(x_0) = 0$, of the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$.*

Proof of Lemma 11. Suppose the outcome $(\lambda_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$ is a pairwise stable outcome of the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$ with $\boldsymbol{u}(x_0) = 0$ and let $(\boldsymbol{\lambda}, \boldsymbol{u}, \boldsymbol{v})$ be the associated outcome of $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. The measures μ_0 and ν_0 have been constructed so that $\lambda_0(x_0, y_0) = 1 + \lambda_0(X \times Y) > 0$ holds for any full match λ_0 in the augmented matching model. Together with the equality $\boldsymbol{u}(x_0) = 0$, the feasibility condition (24) for types (x_0, y_0) in the augmented matching model then implies $\boldsymbol{v}_0(y_0) = 0$. For any type $x \in \text{supp}(\mu)$, (29) in the augmented matching model then implies $\boldsymbol{u}(x) \geq \phi_0(x, y_0, 0) = \underline{\boldsymbol{u}}(x)$ and similarly $\boldsymbol{v}(y) \geq \psi_0(y, x_0, 0) = \underline{\boldsymbol{v}}(y)$ for all $y \in \text{supp}(\nu)$. Thus, the participation constraints (27)–(28) in the associated outcome $(\lambda, \boldsymbol{u}, \boldsymbol{y})$ for the matching model hold. Next, the incentive constraints (29) in the augmented matching model,

$$\boldsymbol{u}_0(x) \geq \phi_0(x, y, \boldsymbol{v}(y)) \quad \forall (x, y) \in \text{supp}(\nu_0) \times \text{supp}(\mu_0),$$

imply

$$\boldsymbol{u}(x) \geq \phi(x, y, \boldsymbol{v}(y)) \quad \forall (x, y) \in \text{supp}(\nu) \times \text{supp}(\mu),$$

which are the incentive constraints in the matching model. It remains to check the feasibility conditions (24)–(26) to infer that $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a stable outcome of $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. As λ and λ_0 coincide on $X \times Y$, the feasibility conditions for the augmented matching model immediately imply $\boldsymbol{u}(x) = \phi(x, y, \boldsymbol{v}(y))$ for all (x, y) in the support of λ, which is (24). It remains only to show that buyers x in the support of $\mu - \lambda_X$ and sellers y in the support of

$\mu - \lambda_Y$ receive their reservation utilities. For such types, we have that (x, y_0) and (y, x_0) are in the support of λ_0, so that (recalling the equalities $\boldsymbol{u}_0(x_0) = \boldsymbol{v}_0(y_0) = 0$ and the definition of ϕ_0), the feasibility condition

$$\boldsymbol{u}_0(x) = \phi(x, y, \boldsymbol{v}_0(y)), \quad \forall (x, y) \in \operatorname{supp}(\lambda_0)$$

for the augmented matching model imply $\boldsymbol{u}(x) = \underline{\boldsymbol{u}}(x)$ and $\boldsymbol{v}(y) = \underline{\boldsymbol{v}}(y)$, which is the desired result.

Conversely, suppose the outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a stable outcome of the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. Let the profiles $\boldsymbol{u}_0 \in \boldsymbol{B}(X_0)$ and $\boldsymbol{u}_0 \in \boldsymbol{B}(X_0)$ agree with $\boldsymbol{u}$ amd $\boldsymbol{v}$ on X and Y and satisfy $\boldsymbol{u}_0(x_0) = 0$ and $\boldsymbol{v}_0(x_0) = 0$. It suffices to show that $(\lambda_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$ is a pairwise stable outcome of the matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$. The equalities $\boldsymbol{u}_0(x_0) = 0$ and $\boldsymbol{v}_0(y_0) = 0$ hold by construction. Feasibility and the conditions for pairwise stability follow from the feasibility and stability conditions for $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ in the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ via arguments analogous to those establishing the previous direction. ■

This completes the proof of Corollary 5. ■

A.13 Proof of Proposition 7.

Let $(\lambda_1, \boldsymbol{u}_1, \boldsymbol{v}_1)$ and $(\lambda_2, \boldsymbol{u}_2, \boldsymbol{v}_2)$ be pairwise stable outcomes. Because the type measures μ and ν have full support (Assumption 2), Proposition 5.3 then implies that $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ as well as $\boldsymbol{u}_2$ and $\boldsymbol{v}_2$ implement each other.

To show that $\mathbb{U}$ and $\mathbb{V}$ are sublattices of $\boldsymbol{B}(X)$ and $\boldsymbol{B}(Y)$, it suffices to show that there exist full matches λ_3 and λ_4 such that $(\lambda_3, \boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2)$ and $(\lambda_4, \boldsymbol{u}_1 \wedge \boldsymbol{u}_2, \boldsymbol{v}_1 \vee \boldsymbol{v}_2)$ are pairwise stable outcomes. The conditions for the pairwise stability of these two outcomes differ from each other only by a reversal of the role of the buyer profiles and the seller profiles, so that we may focus on the first of these, namely the existence of a full match λ_3 such that $(\lambda_3, \boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2)$ is a pairwise stable outcome.

Because $\boldsymbol{v}_1$ implements $\boldsymbol{u}_1$ and $\boldsymbol{v}_2$ implements $\boldsymbol{u}_2$, it is immediate from the fact that the implementation maps are dualities that $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$ implements $\boldsymbol{u}_1 \vee \boldsymbol{u}_2$ (cf. Corollary 1.4 and the discussion at the beginning of Section 3.4.2). Hence, from Proposition 5.1 and 5.2 it suffices to construct a full match λ_3 with $\operatorname{supp}(\lambda_3) \subseteq \Gamma_{\boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2}$ to obtain the desired pairwise stable outcome $(\lambda_3, \boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2)$.

To simplify notation throughout the following, let $\boldsymbol{u}_3 = \boldsymbol{u}_1 \vee \boldsymbol{u}_2$ and $\boldsymbol{v}_3 = \boldsymbol{v}_1 \wedge \boldsymbol{v}_2$. Using this notation, we may rewrite the condition $\operatorname{supp}(\lambda_3) \subseteq \Gamma_{\boldsymbol{u}_1 \vee \boldsymbol{u}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_2}$ as

$$(x, y) \in \operatorname{supp}(\lambda_3) \Rightarrow \boldsymbol{u}_3(x) = \phi(x, y, \boldsymbol{v}_3(y)). \tag{56}$$

Our task is to construct a full match λ_3 satisfying (56). To do so, we define

$$Y_1 = \{y \in Y : \boldsymbol{v}_1(y) < \boldsymbol{v}_2(y)\}$$

and

$$X_1 = \{x \in X : \boldsymbol{Y}_{\boldsymbol{v}_2}(x) \cap Y_1 \neq \emptyset\}.$$

Let $X_2 = X \setminus X_1$ and $Y_2 = Y \setminus Y_1$ denote the complements of X_1 and Y_1.

*S*tep 1: The sets X_1, X_2, Y_1, and Y_2 are measurable.

That $Y_1 \subseteq Y$ is measurable is immediate from the continuity of the implementable assignments $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, which ensures that Y_1 is open in Y. The argmax correspondence $\boldsymbol{Y}_{\boldsymbol{v}_2}$ is weakly measurable (cf. footnote 39), and hence the pre-image of the open set Y_1 under $\boldsymbol{Y}_{\boldsymbol{v}_2}$, namely X_1, is measurable. As the complements of measurable sets, X_2 and Y_2 are measurable.

*S*tep 2: The measures λ_1 and λ_2 are both concentrated on $(X_1 \times Y_1) \cup (X_2 \times Y_2)$.

Recall that $\boldsymbol{v}_2$ and $\boldsymbol{u}_2$ implement each other. By definition of X_1 and Lemma 2, we thus have that $\Gamma_{\boldsymbol{u}_2,\boldsymbol{v}_2}$ and $X_2 \times Y_1$ do not intersect each other. Because $\mathrm{supp}(\lambda_2)$ is contained in $\Gamma_{\boldsymbol{u}_2,\boldsymbol{v}_2}$ (Proposition 5.1) it follows that the support of λ_2 does not intersect $X_2 \times Y_1$ so that

$$\lambda_2(X_2 \times Y_1) = 0 \tag{57}$$

holds. Because λ_2 is a full match, (57) implies $\lambda_2(X_1 \times Y_1) = \nu(Y_1)$. Consequently, we have

$$\mu(X_1) \geq \lambda_2(X_1 \times Y_1) = \nu(Y_1), \tag{58}$$

where the inequality obtains because λ_2 is a match.

Next, we have

$$\lambda_1(X_1 \times Y_2) = 0. \tag{59}$$

To establish this, consider any $x' \in X_1$. By definition of X_1, there exists $y' \in Y_1$ such that $\boldsymbol{u}_2(x') = \phi(x', y', \boldsymbol{v}_2(y')) \geq \phi(x', y, \boldsymbol{v}_2(y))$, with the inequality holding for all $y \in Y$. As $\boldsymbol{v}_1(y') < \boldsymbol{v}_2(y')$ holds (because $y' \in Y_1$) and $\boldsymbol{v}_1$ implements $\boldsymbol{u}_1$ we obtain

$$\boldsymbol{u}_1(x') \geq \phi(x', y', \boldsymbol{v}_1(y')) > \phi(x', y', \boldsymbol{v}_2(y')) \geq \phi(x', y, \boldsymbol{v}_2(y))$$

for all $y \in Y$. As $\boldsymbol{v}_1(y) \geq \boldsymbol{v}_2(y)$ holds for all $y \in Y_2$ this implies $\boldsymbol{u}_1(x') > \phi(x', y, \boldsymbol{v}_1(y))$ for all $y \in Y_2$. As $(\lambda_1, \boldsymbol{u}_1, \boldsymbol{v}_1)$ is pairwise stable, this implies that there does not exist $(x, y) \in X_1 \times Y_2$ contained in the support of λ_1, establishing (59).

Because λ_1 is a match, we have $\nu(Y_1) \geq \lambda_1(X_1 \times Y_1)$. Using the fact that λ_1 is a full match, (59) implies $\lambda_1(X_1 \times Y_1) = \mu(X_1)$, and hence we have

$$\nu(Y_1) \geq \lambda_1(X_1 \times Y_1) = \mu(X_1). \tag{60}$$

Combining (58) and (60) yields

$$\lambda_1(X_1 \times Y_1) = \lambda_2(X_1 \times Y_1) = \mu(X_1) = \nu(Y_1). \tag{61}$$

Because λ_1 and λ_2 are matches, this in turn implies $\lambda_1(X_2 \times Y_1) = 0$ and $\lambda_2(X_1 \times Y_2) = 0$, finishing the argument for this step.

*S*tep 3: Completion of the proof that $\mathbb{U}$ and $\mathbb{V}$ are sublattices.

By Step 1, setting

$$\lambda_3(\tilde{X} \times \tilde{Y}) = \lambda_1((\tilde{X} \cap X_1) \times (\tilde{Y} \cap Y_1)) + \lambda_2((\tilde{X} \cap X_2) \times (\tilde{Y} \cap Y_2)) \tag{62}$$

for all measurable $\tilde{Y} \subseteq Y$ and $\tilde{X} \subseteq X$ defines a measure on $X \times Y$. By Step 2, λ_3 is a full match. It remains to show (56). To obtain this we show first that $\boldsymbol{u}_3(x) = \phi(x, y, \boldsymbol{v}_3(y)$

holds on a subset of $X \times Y$ on which λ_3 is concentrated and then use a continuity argument to extend the result to the support of λ_3.

By construction, λ_3 is concentrated on $(X_1 \times Y_1) \cup (X_2 \times Y_2)$. It is therefore also concentrated on the union of $\text{supp}(\lambda_3) \cap (X_1 \times Y_1)$ and $\text{supp}(\lambda_3) \cap (\mathcal{X} \times Y_2)$, where $\mathcal{X}$ is any measurable subset of X_2 satisfying $\lambda_3(\mathcal{X} \times Y_2) = \lambda_3(X_2 \times Y_2)$.

Consider $(x', y') \in \text{supp}(\lambda_3) \cap (X_1 \times Y_1)$. By construction of λ_3 we then have $(x', y') \in \text{supp}(\lambda_1)$, implying $\boldsymbol{u}_1(x') = \phi(x', y', \boldsymbol{v}_1(y'))$. As $y' \in Y_1$, we have $\boldsymbol{v}_1(y') = \boldsymbol{v}_3(y')$. As $x' \in X_1$, the argument that we have used to establish (59) in Step 2 yields $\boldsymbol{u}_1(x') > \boldsymbol{u}_2(x')$ and thus $\boldsymbol{u}_3(x') = \boldsymbol{u}_1(x')$, establishing (56) for the case under consideration.

Let

$$\mathcal{X} = \{x \in X_2 \mid \boldsymbol{Y}_{\boldsymbol{v}_1} \cap Y_2 \neq \emptyset\}$$

An argument akin to the one used in Step 1 of the proof shows that $\mathcal{X}$ is measurable.[40] By definition of $\mathcal{X}$, $(x, y) \in (X_2 \setminus \mathcal{X}) \times Y_2$ implies $(x, y) \notin \text{supp}(\lambda_1)$, so that $\lambda_1((X_2 \setminus \mathcal{X}) \times Y_2) = 0$ holds. Because λ_1 is a full match, this in turn implies $\lambda_1((X_2 \setminus \mathcal{X}) \times Y_1) = \mu(X_2 \setminus \mathcal{X})$ with $\lambda_1(X_2 \times Y_1) = 0$ (cf. Step 2 of the proof) then implying $\mu(X_2 \setminus \mathcal{X}) = 0$, yielding $\mu(\mathcal{X}) = \mu(X_2)$. As $\lambda_3(\mathcal{X} \times Y_2) = \mu(\mathcal{X})$ and $\lambda_3(X_2 \times Y_2) = \mu(X_2)$ holds, this establishes the requisite property $\lambda_3(\mathcal{X} \times Y_2) = \lambda_3(X_2 \times Y_2)$ of the set $\mathcal{X}$.

Consider now $(x', y') \in \text{supp}(\lambda_3) \cap (\mathcal{X} \times Y_2)$. By construction of λ_3 we then have $(x', y') \in \text{supp}(\lambda_2)$, implying $\boldsymbol{u}_2(x') = \phi(x', y', \boldsymbol{v}_2(y'))$. As $y' \in Y_2$, we have $\boldsymbol{v}_3(y') = \boldsymbol{v}_2(y')$, so that it remains to establish $\boldsymbol{u}_2(x') \geq \boldsymbol{u}_1(x')$ to obtain (56) for the case under consideration. Suppose to the contrary that $\boldsymbol{u}_1(x') > \boldsymbol{u}_2(x')$ holds. As $\boldsymbol{v}_2(y) \leq \boldsymbol{v}_1(y)$ holds on Y_2 this implies $\boldsymbol{u}_1(x') > \phi(x', y, \boldsymbol{v}_1(y))$ for all $y \in Y_2$, which contradicts $x' \in \mathcal{X}$.

Finally, consider any $(x', y') \in \text{supp}(\lambda_3)$. As λ_3 is concentrated on the union of $\text{supp}(\lambda_3) \cap (X_1 \times Y_1)$ and $\text{supp}(\lambda_3) \cap (\mathcal{X} \times Y_2)$, there then exists a sequence $(x_n, y_n)_{n=1}^{\infty}$ in this union which converges to (x', y'). As shown above $\boldsymbol{u}_3(x_n) = \phi(x_n, y_n, \boldsymbol{v}_3(y_n))$ holds for all n in this sequence. As ϕ, $\boldsymbol{v}_3$ and $\boldsymbol{u}_3$ are all continuous, the convergence of $(x_n, y_n)_{n=1}^{\infty}$ to (x', y') implies $\boldsymbol{u}_3(x') = \phi(x', y', \boldsymbol{v}_3(y'))$, which is the desired result.

It remains to show that the set of pairwise stable outcomes for the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ is closed. Let $(\lambda_k, \boldsymbol{u}_k, \boldsymbol{v}_k)$ be a sequence of pairwise stable outcomes for the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ converging to $(\lambda^*, \boldsymbol{u}^*, \boldsymbol{v}^*)$. Using the assumption that $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ have full support, Proposition 5 implies that $(\boldsymbol{u}_k, \boldsymbol{v}_k)$ implement each other for all k. The same arguments as in the proof of Lemma 10 (in Appendix A.11) then imply that $(\lambda^*, \boldsymbol{u}^*, \boldsymbol{v}^*)$ is a pairwise stable outcome for $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$. ■

A.14 Proof of Proposition 8

We establish that the set of stable buyer profiles of the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$, denoted by $\mathbb{U}_s$ in the following, is a complete sublattice of $\boldsymbol{B}(X)$; the argument for the case of stable seller profiles is analogous.

From Lemma 11 in the proof of Proposition 5 (Appendix A.11) an outcome $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is stable in the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ if and only if the associated full outcome

[40] As the complement of the open set Y_1, the set Y_2 is closed with Theorem 17.20 in Aliprantis and Border (2006) then ensuring that $\{x \in X \mid \boldsymbol{Y}_{\boldsymbol{v}_1} \cap Y_2 \neq \emptyset\}$ is measurable. As the intersection of this set with the measurable set X_2, the set $\mathcal{X}$ is measurable.

$(\lambda_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$ is a pairwise stable outcome satisfying the initial condition $\boldsymbol{u}_0(x_0) = 0$ in the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\boldsymbol{u}}_0, \underline{\boldsymbol{v}}_0)$. Denote the set of pairwise stable buyer profiles satisfying the initial condition $\boldsymbol{u}_0(x_0) = 0$ in the augmented matching model by $\mathbb{U}_a$. With the obvious notational convention for the profile $(\boldsymbol{u}_0(x_0), \boldsymbol{u})$ of the augmented matching model, we then have $(\boldsymbol{u}_0(x_0), \boldsymbol{u}) \in \mathbb{U}_a$ if and only if both $\boldsymbol{u}_0(x_0) = 0$ and $\boldsymbol{u} \in \mathbb{U}_s$ hold. It is then immediate that $\mathbb{U}_s$ is a complete sublattice of $\boldsymbol{B}(X)$ if $\mathbb{U}_a$ is a complete sublattice of $\boldsymbol{B}(X_0)$.

To show that $\mathbb{U}_a$, which is non-empty by Proposition 5, is a complete sublattice of $\boldsymbol{B}(X_0)$, we first observe that $\mathbb{U}_a$ is the intersection of two closed sublattices of $\boldsymbol{B}(X_0)$, namely the set of pairwise stable buyer profiles of the augmented matching model (which is closed by Proposition 6 and a sublattice by Proposition 7) and the set of profiles $\boldsymbol{u}_0 \in \boldsymbol{B}(X_0)$ satisfying $\boldsymbol{u}_0(x_0) = 0$ (which is obviously a sublattice and closed). Hence, $\mathbb{U}_a$ is a closed sublattice of $\boldsymbol{B}(X_0)$. Further, the closed sublattice $\mathbb{U}_a$ is bounded, with the profile $\underline{\boldsymbol{u}}_0$ providing a lower bound and the profile $\Phi\underline{\boldsymbol{v}}_0$ providing an upper bound. Hence (Corollary 4.3), $\mathbb{U}_a$ is a compact sublattice and therefore (by the same argument as in the proof of Lemma 4, cf. Footnote 38) complete. ■

A.15 Proof of Lemma 6

Step 1: We first argue that it is without loss of generality to restrict the principal's choice set to implementable tariffs: Let $(\lambda, \boldsymbol{u}, \boldsymbol{v}) \in \mathbb{M} \times \boldsymbol{B}(X) \times \boldsymbol{B}(Y)$ be any triple satisfying the constraints in the principal's maximization problem defined in Section 5.1. Consider the triple $(\lambda, \boldsymbol{u}, \Psi\boldsymbol{u})$. The tariff $\Psi\boldsymbol{u}$ is implementable and implements $\boldsymbol{u}$ and, further, implements any selection from $\boldsymbol{Y_v}$ (Corollary 3.1), so that $\boldsymbol{Y_v}(x) \subseteq \boldsymbol{Y}_{\Psi\boldsymbol{u}}(x)$ holds for all $x \in X$. Consequently, we have $\Gamma_{\boldsymbol{u},\boldsymbol{v}} \subseteq \Gamma_{\boldsymbol{u},\Psi\boldsymbol{u}}$, ensuring that the triple $(\lambda, \boldsymbol{u}, \Psi\boldsymbol{u})$ is in the feasible set of the principal's problem. As we have noted in the text following equation (34), the feasibility of $(\lambda, \boldsymbol{u}, \Psi\boldsymbol{u})$ implies that it results in the same expected payoff as $(\lambda, \boldsymbol{u}, \boldsymbol{v})$.

Step 2: From Step 1 we can restrict attention to $(\lambda, \boldsymbol{u}, \boldsymbol{v}) \in \mathbb{M} \times \boldsymbol{B}(X) \times \boldsymbol{I}(Y)$ when considering the principal's problem. As $\boldsymbol{v} \in \boldsymbol{I}(Y)$ implements $\boldsymbol{u} \in \boldsymbol{B}(Y)$ if and only if $\boldsymbol{u} = \Phi\boldsymbol{v}$, we can eliminate the first constraint from the principal's problem and substitute this equality in the remaining constraints. The resulting problem is:

$$\max_{\boldsymbol{v}\in\boldsymbol{I}(Y),\lambda\in\mathbb{M}} \int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(x, y)$$
$$\text{s.t. } \operatorname{supp}(\lambda) \subseteq \Gamma_{\Phi\boldsymbol{v},\boldsymbol{v}} \text{ and } \Phi\boldsymbol{v} \geq \underline{\boldsymbol{u}}.$$

Because implementable profiles are continuous (Proposition 2), the objective function in this problem is well-defined for all $\boldsymbol{v} \in \boldsymbol{I}(Y)$ and $\lambda \in \mathbb{M}$. Using (i) the definition of $F(\boldsymbol{v}, \lambda)$ in (35), (ii) observing that the first constraint is equivalent to $\lambda \in G(\boldsymbol{v})$, where $G(\boldsymbol{v})$ is defined in (36), and (iii) using the order reversal property of the implementation maps (Corollary 1.2) to transform the second constraint into $\boldsymbol{v} \leq \Psi\underline{\boldsymbol{u}}$, we may rewrite the above problem as

$$\max_{\{\boldsymbol{v}\in\boldsymbol{I}(Y):\boldsymbol{v}\leq\Psi\underline{\boldsymbol{u}}\}} \left[\max_{\lambda\in G(\boldsymbol{v})} F(\boldsymbol{v}, \lambda)\right].$$

Step 3: Let $(\boldsymbol{v}_n)_{n=1}^{\infty}$ converge in norm to $\boldsymbol{v}$ and let $(\lambda_n)_{n=1}^{\infty}$ converge weakly to λ. Then for any $\varepsilon > 0$, we can find N such that for all $n \geq N$, we have

$$\begin{aligned}
F(\boldsymbol{v},\lambda) - 2\varepsilon &= \int_X \int_Y \pi(x,y,\boldsymbol{v}(y)) d\lambda(x,y) - 2\varepsilon \\
&\leq \int_X \int_Y \pi(x,y,\boldsymbol{v}(y)) d\lambda_n(x,y) - \varepsilon \\
&= \int_X \int_Y \pi(x,y,\boldsymbol{v}(y) - \varepsilon) d\lambda_n(x,y) \\
&\leq \int_X \int_Y \pi(x,y,\boldsymbol{v}_n(y)) d\lambda_n(x,y) \\
&\leq \int_X \int_Y \pi(x,y,\boldsymbol{v}(y) + \varepsilon) d\lambda_n(x,y) \\
&= \int_X \int_Y \pi(x,y,\boldsymbol{v}(y)) d\lambda_n(x,y) + \varepsilon \\
&\leq \int_X \int_Y \pi(x,y,\boldsymbol{v}(y)) d\lambda(x,y) + 2\varepsilon \\
&= F(\boldsymbol{v},\lambda) + 2\varepsilon.
\end{aligned}$$

The two central inequalities follow from the convergence of $(\boldsymbol{v}^n)_{n=1}^{\infty}$, and the two remaining inequalities from the convergence of $(\lambda_n)_{n=1}^{\infty}$. Combining the middle and outside two terms, we have $F(\boldsymbol{v},\lambda) - 2\varepsilon \leq F(\boldsymbol{v}_n,\lambda_n) \leq F(\boldsymbol{v},\lambda) + 2\varepsilon$. Hence, the function $F(\boldsymbol{v},\lambda)$ is continuous.

Step 4: For $\boldsymbol{v} \in \boldsymbol{I}(Y)$, the correspondence $G(\boldsymbol{v})$ defined in (36) is non-empty and compact valued and upper hemicontinuous. To show that $G(\boldsymbol{v})$ is non-empty valued, let $\boldsymbol{y}$ be a measurable selection (cf. footnote 39) from $\boldsymbol{Y}_{\boldsymbol{v}}$ and let $\lambda_{\boldsymbol{y}}$ be the associated deterministic measure (cf. (32)). As $\boldsymbol{v}$ and $\Phi\boldsymbol{v}$ implement each other, the same argument as in the first paragraph of the proof of Lemma 5 yields that the support of $\lambda_{\boldsymbol{y}}$ is contained in $\Gamma_{\Phi\boldsymbol{v},\boldsymbol{v}}$. Hence, $G(\boldsymbol{v})$ is non-empty valued.

To obtain the other two properties, define the function $H : X \times Y \times \boldsymbol{I}(Y) \to \mathbb{R}$ by $H(x,y,\boldsymbol{v}) = \phi(x,y,\boldsymbol{v}(y)) - \Phi\boldsymbol{v}(x)$. Notice that H is continuous because ϕ and Φ are (Lemma 1). In addition, $H(x,y,\boldsymbol{v}) \leq 0$, with equality if and only if $(x,y) \in \Gamma_{\Phi\boldsymbol{v},\boldsymbol{v}}$. Now consider the maximization problem $\max_{\lambda\in\mathbb{M}} \hat{H}(\boldsymbol{v},\lambda)$, where $\hat{H} : \boldsymbol{I}(Y) \times \mathbb{M} \to \mathbb{R}$ is defined by $\hat{H}(\boldsymbol{v},\lambda) = \int_X \int_Y H(x,y,\boldsymbol{v}) d\lambda(x,y)$. For any $\boldsymbol{v}$, we have $\hat{H}(\boldsymbol{v},\lambda) \leq 0$, with equality if and only if $\text{supp}(\lambda) \in \Gamma_{\Phi\boldsymbol{v},\boldsymbol{v}}$. The argmax correspondence for this maximization problem thus is $G(\boldsymbol{v})$. We have noted that $H(x,y,\boldsymbol{v})$ is continuous and hence so is $\hat{H}(\boldsymbol{v},\lambda)$, and $\mathbb{M}$ is compact by Prokhorov's theorem (Shiryaev, 1996, p. 318). An application of Berge's maximum theorem (Ok, 2007, p. 306) then ensures that $G(\boldsymbol{v})$ is compact-valued and upper hemicontinuous.

Step 5: Fix $\boldsymbol{v} \in \boldsymbol{I}(Y)$ and consider the problem appearing in (37):

$$\Pi(\boldsymbol{v}) = \max_{\lambda\in G(\boldsymbol{v})} F(\boldsymbol{v},\lambda).$$

We have shown in Step 3 that $F(\boldsymbol{v},\lambda)$ is continuous and in Step 4 that $G(\boldsymbol{v})$ is nonempty and compact valued. Therefore, Weierstrass' extreme value theorem ensures that this problem has a solution so that the function $\Pi : \boldsymbol{I}(Y) \to \mathbb{R}$ is well-defined. Further, because the correspondence G is also upper hemicontinuous (Step 4), Berge's maximum theorem (Ok, 2007, p. 306) ensures that Π is upper semicontinuous.

Step 6: Let v^* solve the problem

$$\max_{\{\boldsymbol{v}\in\boldsymbol{I}(Y):\ \boldsymbol{v}\leq\Psi\underline{\boldsymbol{u}}\}} \Pi(\boldsymbol{v})$$

and let λ^* be an element of $\arg\max_{\lambda\in G(\boldsymbol{v}^*)} F(\boldsymbol{v}^*, \lambda)$. Then it is immediate from (37) that (v^*, λ^*) solves the problem

$$\max_{\{\boldsymbol{v}\in\boldsymbol{I}(Y):\boldsymbol{v}\leq\Psi\underline{\boldsymbol{u}}\}} \left[\max_{\lambda\in G(\boldsymbol{v})} F(\boldsymbol{v}, \lambda)\right].$$

As noted in Step 2, this implies that $(\lambda^*, \Phi\boldsymbol{v}^*, \boldsymbol{v}^*)$ solves the principal's problem when the principal is restricted to $\boldsymbol{v} \in \boldsymbol{I}(Y)$. Step 1 then ensures that the triple $(\lambda^*, \Phi\boldsymbol{v}^*, \boldsymbol{v}^*)$ solves the principal's problem. ∎

A.16 Proof of Proposition 9

We proceed in two steps, first establishing the existence of a solution $\boldsymbol{v}$ to the nonlinear pricing problem (38) and then showing that in the associated solution $(\lambda, \Phi\boldsymbol{v}, \boldsymbol{v})$ to the principal's problem, the measure λ can be taken to be deterministic.

Step 1: We first show that we can restrict attention to a bounded set of tariffs. To simplify notation, let $\overline{\boldsymbol{v}} = \Psi\underline{\boldsymbol{u}}$ denote the upper bound for the feasible set in the nonlinear pricing problem. By Proposition 2, we have $\overline{\boldsymbol{v}} \in \boldsymbol{I}(Y)$, so that $\Pi(\overline{\boldsymbol{v}})$ is well-defined. To obtain a lower bound, let $v^\dagger \in \mathbb{R}$ be such that for all $(x, y) \in X \times Y$

$$\pi(x, y, v^\dagger) < \Pi(\overline{\boldsymbol{v}}).$$

The existence of such a $v^\dagger$ is ensured because π satisfies the full range condition in Assumption 3 and X and Y are compact. By Assumption 1, there also exists $\underline{v} \in \mathbb{R}$ such that, for all (x, y) in $X \times Y$ and $v \leq \underline{v}$, we have

$$\phi(x, y, v) > \max_{\hat{y}\in Y} \phi(x, \hat{y}, v^\dagger).$$

The second of these displayed inequalities ensures that for any tariff $\boldsymbol{v} \in \boldsymbol{I}(Y)$ with the property that $\boldsymbol{v}(y) \leq \underline{v}$ holds for some $y \in Y$, we have that $(\hat{x}, \hat{y}) \in \Gamma_{\Phi\boldsymbol{v},\boldsymbol{v}}$ implies $\boldsymbol{v}(\hat{y}) < v^\dagger$. From the first displayed inequality, this ensures that $F(\boldsymbol{v}, \lambda) < \Pi(\overline{\boldsymbol{v}})$ holds for all $\lambda \in G(\boldsymbol{v})$, implying that $\Pi(\boldsymbol{v}) < \Pi(\overline{\boldsymbol{v}})$ holds for any such tariff. Hence, $\Pi(\boldsymbol{v}) \geq \Pi(\overline{\boldsymbol{v}})$ implies $\boldsymbol{v}(y) \geq \underline{v}$ for all $y \in Y$ and there thus exists a tariff $\underline{\boldsymbol{v}} \in \boldsymbol{I}(Y)$ such that $\Pi(\boldsymbol{v}) \geq \Pi(\overline{\boldsymbol{v}})$ implies $\boldsymbol{v} \geq \underline{\boldsymbol{v}}$.

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 4.1), 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 4.3 $\mathcal{V}$ is therefore compact. As $\overline{\boldsymbol{v}}$ is an element of both $\mathcal{V}$ and $\boldsymbol{I}(Y)$ this set is also non-empty.

As Π is upper semicontinuous (Lemma 6), Weierstrass' extreme value theorem for upper semicontinuous functions (Ok, 2007, p.234) then implies that the problem

$$\max_{\{\boldsymbol{v}\in\boldsymbol{I}(Y):\ \underline{\boldsymbol{v}}\leq\boldsymbol{v}\leq\overline{\boldsymbol{v}}\}} \Pi(\boldsymbol{v})$$

has a solution $\boldsymbol{v}^*$. We obviously have $\Pi(\boldsymbol{v}^*) \geq \Pi(\overline{\boldsymbol{v}})$ and hence $\Pi(\boldsymbol{v}^*) \geq \Pi(\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{I}(Y)$ satisfying (using the definition of $\underline{\boldsymbol{v}}$) $\boldsymbol{v} \leq \overline{\boldsymbol{v}} = \Psi\underline{\boldsymbol{u}}$, ensuring that $\boldsymbol{v}^*$ solves the nonlinear pricing problem (38).

Step 2: Let $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ be feasible in the principal's problem with $\boldsymbol{v} \in \boldsymbol{I}(Y)$. We first observe that $\max_{y\in\boldsymbol{Y}_{\boldsymbol{v}}(x)} \pi(x, y, \boldsymbol{v}(y))$ is a measurable function of x and that there exists a measurable assignment $\boldsymbol{y}^*$ solving this maximization problem for all x. This follows from Aliprantis and Border (2006, Theorem 18.19) upon observing that (i) the function $(x, y) \to \pi(x, y, \boldsymbol{v}(y))$ is continuous on its domain $X \times Y$ (from Proposition 2 and Assumption 3) and thus a Caratheodory function and (ii) the properties of the correspondence $\boldsymbol{Y}_{\boldsymbol{v}}$ noted in Corollary 2 imply that this correspondence has a closed graph, ensuring that it is weakly measurable (Aliprantis and Border, 2006, Theorem 18.20 and Lemma 18.2).

We can then write

$$\begin{aligned}
F(\boldsymbol{v}, \lambda) &= \int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(x, y) \\
&= \int_X \left(\int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(y \mid x) \right) d\mu(x) \\
&\leq \int_{x\in X} \max_{y\in\boldsymbol{Y}_{\boldsymbol{v}}(x)} \pi(x, y, \boldsymbol{v}(y)) d\mu(x) \\
&= \int_{x\in X} \pi(x, \boldsymbol{y}^*(x), \boldsymbol{v}(\boldsymbol{y}^*(x))) d\mu(x) \\
&= F(\boldsymbol{v}, \lambda_{\boldsymbol{y}^*}),
\end{aligned}$$

where the equality in the second line follows from the disintegration theorem (Chang and Pollard, 1997, Theorem 1), with $\lambda(\cdot \mid x)$ being the disintegration measure on $\{x\} \times Y$ for each $x \in X$. The inequality holds because the support of $\lambda(\cdot \mid x)$ is contained in $\boldsymbol{Y}_{\boldsymbol{v}}(x)$ for μ-almost all $x \in X$. The equality on the penultimate line is by definition of $\boldsymbol{y}^*$. As $(\lambda_{\boldsymbol{y}^*}, \boldsymbol{u}, \boldsymbol{v})$ is feasible in the principal's problem and this problem has a solution, the inequality $F(\boldsymbol{v}, \lambda) \leq F(\boldsymbol{v}, \lambda_{\boldsymbol{y}^*})$ implies that the principal's problem has a deterministic solution. ∎

A.17 Proof of Proposition 10

Suppose $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ solves the principal's problem with $\boldsymbol{u}(x) > \underline{\boldsymbol{u}}(x)$ for all $x \in X$. From Proposition 9 there exists a deterministic match $\lambda_{\boldsymbol{y}}$, such that $(\lambda_{\boldsymbol{y}}, \boldsymbol{u}, \boldsymbol{v})$ is also a solution to the principal's problem. By the same argument as the one proving Lemma 5, we can take $\boldsymbol{y}$ to be implementable and therefore (by assumption) to be strongly implementable. From Lemma 4 there thus exists a profile $\boldsymbol{u}^*$ such that $(\boldsymbol{u}^*, \boldsymbol{y})$ is implementable, $\boldsymbol{u} \geq \boldsymbol{u}^* \geq \underline{\boldsymbol{u}}$ holds, and there exists $x \in X$ such that $\boldsymbol{u}(x) > \boldsymbol{u}^*(x)$ for some $x \in X$. As both $\boldsymbol{u}$ and $\boldsymbol{u}^*$ are implementable (and therefore continuous by Proposition 2) the set $\mathcal{X} = \{x \in X \mid u(x) > u^*(x)\}$ is measurable. Because μ has full support, we have $\mu(\mathcal{X}) > 0$.

Now, let $\boldsymbol{v}^* = \Psi \boldsymbol{u}^*$. Then $\boldsymbol{v}^*$ implements $(\boldsymbol{u}^*, \boldsymbol{y})$ (Corollary 3.1) and the triple $(\boldsymbol{y}, \boldsymbol{u}^*, \boldsymbol{v}^*)$ is therefore feasible in the principal's problem We also have that the principal obtains a strictly higher expected payoff from $(\boldsymbol{y}, \boldsymbol{u}^*, \boldsymbol{v}^*)$ than from $(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$, contradicting the optimality of $(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v})$:

$$\begin{aligned} \int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda_{\boldsymbol{y}}(x, y) &= \int_X \int_Y \pi(x, y, \psi(y, x, \boldsymbol{u}(x))) d\lambda_{\boldsymbol{y}}(x, y) \\ < \int_X \int_Y \pi(x, y, \psi(y, x, \boldsymbol{u}^*(x))) d\lambda_{\boldsymbol{y}}(x, y) &= \int_X \int_Y \pi(x, y, \boldsymbol{v}^*(y)) d\lambda_{\boldsymbol{y}}(x, y), \end{aligned}$$

where the equalities follow as in (34) and the strict inequality holds because $\mu(\mathcal{X}) > 0$, ψ is strictly decreasing in its third argument, and π is strictly increasing in its third argument.■

A.18 Proof of Proposition 11

Suppose $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a solution to the principal's problem with $\boldsymbol{u}(x) > \underline{\boldsymbol{u}}(x)$ for all $x \in X$. Then as we have noted in Remark 8, $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a pairwise stable outcome of the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$, where ν is the marginal measure λ_Y of λ on Y and $\underline{\boldsymbol{v}} : Y \to \mathbb{R}$ is an arbitrary continuous function. Let $\mathcal{Y}$ be the support of ν. It exposes the logic of the argument most clearly by first proceeding under the assumption that $\mathcal{Y} = Y$.

The assumption $\mathcal{Y} = Y$ ensures that the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ satisfies Assumption 2, so that this matching model has a pairwise stable outcome $(\hat{\lambda}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$ satisfying $\boldsymbol{u} \geq \hat{\boldsymbol{u}} \geq \underline{\boldsymbol{u}}$, with the first inequality holding strictly for some $x \in X$ (Corollary 7). Because $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{v}}$ implement each other (Proposition 5.3) and the implementation maps are order reversing inverse bijections (cf. (15)), we thus obtain $\boldsymbol{v} \leq \hat{\boldsymbol{v}}$ with strict inequality for some $y \in Y$. From the continuity of the two profiles $\boldsymbol{v}$ and $\hat{\boldsymbol{v}}$ (Proposition 2) and the assumption that ν has full support, we thus obtain

$$\nu(\{y : \boldsymbol{v}(y) < \hat{\boldsymbol{v}}(y)\}) > 0. \tag{63}$$

We can now write

$$\begin{aligned} \int_X \int_Y \pi(x, y, \boldsymbol{v}(y)) d\lambda(x, y) &= \int_Y \hat{\pi}(y, \boldsymbol{v}(y)) d\nu(y) \\ &< \int_Y \hat{\pi}(y, \hat{\boldsymbol{v}}(y)) d\nu(y) \\ &= \int_X \int_Y \pi(x, y, \hat{\boldsymbol{v}}(y)) d\hat{\lambda}(x, y), \end{aligned}$$

where the equalities are from the private-values assumption and the inequality follows from (63) because $\hat{\pi}$ is strictly increasing in its second argument (Assumption 3). We thus obtain $F(\boldsymbol{v}, \lambda) > F(\hat{\boldsymbol{v}}, \hat{\lambda})$. As $(\hat{\lambda}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$ is feasible in the principal's problem, this contradicts the optimality of $(\lambda, \boldsymbol{u}, \boldsymbol{v})$.

If $\mathcal{Y}$ is a strict subset of Y, then the above argument is not directly applicable because the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ violates the full support condition in Assumption 2. It is, however, straightforward to establish a "restriction lemma" (similar in spirit to the extension result of Proposition 5.3) showing that if $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ is a pairwise stable outcome of the matching model $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$, then $(\lambda, \boldsymbol{u}, \boldsymbol{v})$ can be restricted to give a pairwise stable outcome of the matching model derived from $(X, Y, \phi, \mu, \nu, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}})$ by restricting the sets X and Y to the supports $\mathcal{X}$ and $\mathcal{Y}$ of μ and ν. This latter model satisfies Assumption 2, allowing us to repeat the argument above (and in particular to apply Corollary 7). The conclusion of this argument is that the principal can secure a higher payoff than under $(\boldsymbol{\lambda}, \boldsymbol{u}, \boldsymbol{v})$, even if restricted to assigning only decisions in $\mathcal{Y}$ to the agents. ■

References

Adachi, H. (2000): "On a characterization of stable matchings," *Economics Letters*, 68(1), 43–49.

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. (2017): "Duality in Contracting," Discussion paper, University of Melbourne.

Basov, S. (2006): *Multidimensional Screening*. Springer-Verlag.

Becker, G. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813.

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

Bonnet, O., A. Galichon, and M. Shum (2017): "Yogurts Choose Consumers? Identification of Random Utlity Models via Two-Sided Matching," Discussion paper, SSRN.

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.

——— (2003): "Duality and Existence for a Class of Mass Transportation Problems and Economic Applications," in *Advances in Mathematical Economics*, ed. by T. Maruyama, and S. Kusuoka, no. 5, pp. 1–21. Springer-Verlag.

Chade, H., J. Eeckhout, and L. Smith (2017): "Sorting through Search and Matching Models in Economics," *Journal of Economic Literature*, 55(2), 493–544.

Chambers, C., and A. D. Miller (2011): "Rules for Aggregating Preferences," *Social Choice and Welfare*, 36(1), 75–82.

Chang, J. T., and D. Pollard (1997): "Conditioning as Disintegration," *Statistica Neerlandica*, 51(3), 287–317.

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.

Chiappori, P.-A., and B. Salanié (2016): "The Econometrics of Matching Models," *Journal of Economic Literature*, 54(3), 832–861.

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

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

Dhaene, J., M. Denuit, M. J. Goovaerts, R. Kaas, and D. Vyncke (2002): "The Concept of Comonotonicity in Actuarial Science and Finance: Theory," *Insurance: Mathematics and Economics*, 31(1), 133–161.

Dworczak, P., and A. L. Zhang (2017): "Implementability, Walrasian Equilibria, and Efficient Matchings," *Economics Letters*, 153, 57–60.

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

Ferrera, J. (2014): *An Introduction to Nonsmooth Analysis*. Elsevier.

Figalli, A., Y. Kim, and R. McCann (2011): "When is Multidimensional Screening a Convex Program?," *Journal of Economic Theory*, 146(2), 454–478.

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

Gale, D., and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69(1), 9–15.

Galichon, A. (2016): *Optimal Transportation Methods in Economics*. Princeton University Press.

Galichon, A., S. D. Kominers, and S. Weber (2016): "Costly Concessios: An Empirical Framework for Matching with Imperfectly Transferable Utility," SSRN working paper, NYU, Harvard Univesity and Sciences Po.

Glosten, L. R. (1989): "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business*, 62(2), 211–235.

Greinecker, M., and C. Kah (2016): "Pairwise Matching in Large Economies," Discussion paper, University of Innsbruck.

Gretsky, N., J. Ostroy, and W. Zame (1999): "Perfect Competition in the Continuous Assignment Model," *Journal of Economic Theory*, 88(1), 60–118.

Gretsky, N. E., J. M. Ostroy, and W. R. Zame (1992): "The Nonatomic Assignment Model," *Economic Theory*, 2(1), 103–127.

Guesnerie, R., and J. Laffont (1984): “Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm,” *Journal of Public Economics*, 25(3), 329–369.

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

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

Kadan, O., P. J. Reny, and J. Swinkels (2017): “Existence of Optimal Mechanisms in Principal-Agent Problems,” *Econometrica*, 85(3), 769–823.

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 (1986): “The Core of a Game with a Continuum of Players and Finite Coalitions: The Model and some Results,” *Mathematical Social Sciences*, 12(2), 105–137.

——— (1996): “The Nonemptiness of the f-Core of a Game without Side Payments,” *International Journal of Game Theory*, 25(2), 245–258.

Kelley, J. L. (1955): *General Topology*. Springer-Verlag.

Laffont, J.-J., and D. Martimort (2002): *The Theory of Incentives*. Princeton University Press.

Laffont, J.-J., and J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation*. MIT Press.

Lawler, E. L. (2001): *Combinatorial Optimization: Networks and Matroids*. Courier Corporation.

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

Martinez-Legaz, J., and I. Singer (1990): “Dualities between Complete Lattices,” *Optimization*, 21(4), 481–508.

Martinez-Legaz, J., and I. Singer (1995): “Subdifferentials with Respect to Dualities,” *Mathematical Methods of Operations Research*, 42(1), 109–125.

McCann, R. J. (1995): “Existence and Uniqueness of Monotone Measure-Preserving Maps,” *Duke Mathematical Journal*, 80(2), 309–324.

McCann, R. J., and K. S. Zhang (2017): “On Concavity of the Monopolist’s Problem Facing Consumers with Nonlinear Price Preferences,” Discussion paper, University of Toronto.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.

——— (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.

Myerson, R. B. (1979): "Optimal Auction Design," *Mathematics of Operations Research*, 6(1), 58–73.

——— (1982): "Optimal Coordination in Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10(1), 67–81.

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

Nöldeke, G., and L. Samuelson (2015): "Investment and Competitive Matching," *Econometrica*, 83(3), 835–896.

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-Verlag, New York.

Quint, T. (1994): "The Lattice of Core (Sub) Matchings in a Two-sided Matching Market," *Mathematics of Operations Research*, 19(3), 603–617.

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(2), 191–200.

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

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

Shapley, L., and M. Shubik (1972): "The Assignment Game I: The Core," *International Journal of Game Theory*, 1(1), 111–130.

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

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

Stiglitz, J. (1977): "Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market," *Review of Economic Studies*, 44, 407–430.

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*. Cambridge University Press.

Weibull, J. (1989): "A Note on the Continuity of Incentive Schedules," *Journal of Public Economics*, 39(2), 239–243.

Wilson, R. B. (1993): *Nonlinear Pricing*. Oxford University Press.