Econometrica Supplementary Material

SUPPLEMENT TO "THE IMPLEMENTATION DUALITY"
(*Econometrica*, Vol. 86, No. 4, July 2018, 1283–1324)

GEORG NÖLDEKE
Faculty of Business and Economics, University of Basel

LARRY SAMUELSON
Department of Economics, Yale University

APPENDIX B: ARGUMENTS OMITTED FROM THE PAPER

B.1. *Properties of the Inverse Generating Function in Section 2.2*

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.

B.2. *Details for Remark 1*

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

LEMMA B.1: *Let Assumption* 1 *hold*.

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

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

PROOF OF LEMMA B.1: We prove Lemma B.1(1); the proof of Lemma B.1(2) is analogous.

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

Georg Nöldeke: georg.noeldeke@unibas.ch
Larry Samuelson: larry.samuelson@yale.edu
Financial support from National Science Foundation Grant SES-1459158 is gratefully acknowledged.

 https://doi.org/10.3982/ECTA13307

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

$$\mathbf{u}(x) = \phi\big(x, \mathbf{y}(x), \mathbf{t}(x)\big) \geq \phi\big(x, \mathbf{y}(\hat{x}), \mathbf{t}(\hat{x})\big), \tag{B.1}$$

$$\mathbf{t}(x) = \psi\big(\mathbf{y}(x), x, \mathbf{u}(x)\big) \geq \psi\big(\mathbf{y}(x), \hat{x}, \mathbf{u}(\hat{x})\big) \tag{B.2}$$

hold for all $x, \hat{x} \in X$. The equality in (B.2) follows from the equality in (B.1) because ϕ and ψ are inverse and the inequality in (B.2) follows from (B.1) upon reversing the roles of x and $\hat{x}$ in the inequality $\mathbf{u}(x) \geq \phi(x, \mathbf{y}(\hat{x}), \mathbf{t}(\hat{x}))$ and using, again, that ϕ and ψ are inverse.

First, we establish that $\mathbf{u}$ is bounded. Fix $\hat{x} \in X$. The inequality in (B.1) ensures that, for all $x \in X$,

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

where the minimum $\underline{u}$ exists because X is compact and ϕ is continuous. Next, using (B.2), we have

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

for all $x \in X$, where the minimum $\underline{t}$ exists because Y is compact and ψ is continuous. Using the equality in (B.1) and that ϕ is strictly decreasing in its third argument, we then have, for all $x \in X$,

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

where the maximum $\overline{u}$ exists because X and Y are compact and ϕ is continuous. We thus have $\underline{u} \leq \mathbf{u}(x) \leq \overline{u}$ for all $x \in X$, which implies $\mathbf{u} \in \mathbf{B}(X)$. From the equality in (B.2), $\mathbf{t}$ is bounded, too.

Second, we show there exists $\mathbf{v} \in \mathbf{B}(Y)$ implementing $(\mathbf{u}, \mathbf{y})$. 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

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

If there exist $x, \hat{x} \in X$ and $y \in Y$ with $y = \mathbf{y}(x) = \mathbf{y}(\hat{x})$, then the incentive constraints in (B.1) imply $\mathbf{t}(x) = \mathbf{t}(\hat{x})$. Therefore, $\mathbf{v}(y)$ is well-defined for all $y \in Y$ and, because $\mathbf{t}$ is bounded, we have $\mathbf{v} \in \mathbf{B}(Y)$. Finally, using (B.1), it is immediate from the construction of $\mathbf{v}$ that we have

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

for all $(x, y) \in X \times Y$, so that $\mathbf{v}$ implements $(\mathbf{u}, \mathbf{y})$. *Q.E.D.*

B.3. *Details for Remark 2*

To verify that (7) implies the strong implementability of every implementable assignment, we first consider an implementable assignment $\mathbf{y} \in Y^X$. Because $\mathbf{y}$ is implementable, there exists $\mathbf{v} \in \mathbf{B}(Y)$ such that $\mathbf{y}(x) \in \mathbf{Y}_{\mathbf{v}}(x)$ holds for all $x \in X$. Fix any $x_0 \in X$. Because $\mathbf{v}$ implements $\mathbf{y}$, it is immediate that $\mathbf{y}$ is implementable with initial condition (x_0, u_0), where

$u_0 = \phi(x_0, \mathbf{y}(x_0), \mathbf{v}(\mathbf{y}(x_0)))$. Using Assumption 1, for any $t_0 \in \mathbb{R}$ we can find a uniquely determined profile $\hat{\mathbf{v}}$ such that

$$\phi\big(x_0, y, \mathbf{v}(y)\big) - \phi\big(x_0, y, \hat{\mathbf{v}}(y)\big) = t_0, \quad \forall y \in Y. \tag{B.3}$$

The optimal decisions of type x_0 when maximizing against the tariff $\mathbf{v}$ are then identical to the optimal decisions when maximizing against $\hat{\mathbf{v}}$. Further, the same holds for any other type $x_1 \in X$, since (B.3) and (7) ensure that there exists t_1 such that $\phi(x_1, y, \mathbf{v}(y)) - \phi(x_1, y, \hat{\mathbf{v}}(y)) = t_1$ holds for all $y \in Y$. Therefore, if the generating function satisfies (7), then $\mathbf{Y}_{\mathbf{v}}(x) = \mathbf{Y}_{\hat{\mathbf{v}}}(x)$ holds for all $x \in X$, so that $\hat{\mathbf{v}}$ implements $\mathbf{y}$ with initial condition $(x_0, u_0 - t_0)$. As both x_0 and t_0 were arbitrary, this shows that $\mathbf{y}$ is strongly implementable.

Second, consider an implementable assignment $\mathbf{x} \in X^Y$. Then there exists $\mathbf{u} \in \mathbf{B}(X)$ such that $\mathbf{x}(y) \in \mathbf{X}_{\mathbf{u}}(y)$ holds for all $y \in Y$. We first show that, for any $(x_0, t_0) \in X \times \mathbb{R}$, there exists $\hat{\mathbf{u}} \in \mathbf{B}(X)$ satisfying $\mathbf{X}_{\hat{\mathbf{u}}}(y) = \mathbf{X}_{\mathbf{u}}(y)$ for all $y \in Y$ and $\mathbf{u}(x_0) - \hat{\mathbf{u}}(x_0) = t_0$. To do so, we make use of results from Section 3. We may suppose without loss of generality that the profile $\mathbf{u}$ implementing $\mathbf{x}$ is itself implementable (Corollary 4(2)), so that the profile $\mathbf{v}$ implemented by $\mathbf{u}$ also implements $\mathbf{u}$ (Corollary 3(2)). Applying Lemma 2, we then have that the graphs of both $\mathbf{Y}_{\mathbf{v}}$ and $\mathbf{X}_{\mathbf{u}}$ coincide with $\Gamma_{\mathbf{u},\mathbf{v}}$ (with the latter defined in (16)). Now consider $\hat{\mathbf{v}}$ as constructed in the first step above. Using condition (7), we then have $\mathbf{Y}_{\mathbf{v}} = \mathbf{Y}_{\hat{\mathbf{v}}}$. Because $\mathbf{v}$ is implementable, this equality of the argmax correspondences implies that $\hat{\mathbf{v}}$ is also implementable (Remark 6). Applying Corollary 3(1), the profile $\hat{\mathbf{u}}$ implemented by $\hat{\mathbf{v}}$ also implements $\hat{\mathbf{v}}$. Applying Lemma 2 again, it follows that $\mathbf{X}_{\hat{\mathbf{u}}}$ coincides with $\mathbf{X}_{\mathbf{u}}$. Then the equality $\mathbf{u}(x_0) - \hat{\mathbf{u}}(x_0) = t_0$ follows directly from the construction of $\hat{\mathbf{v}}$.

To complete the argument, choose (y_0, v_0) and let $x_0 = \mathbf{x}(y_0)$. Then $\mathbf{u}$ implements $(\mathbf{v}, \mathbf{x})$ with $\mathbf{v}(y_0) = \psi(y_0, x_0, \mathbf{u}(x_0))$. In addition, for any t_0, $\hat{\mathbf{u}}$ implements $(\hat{\mathbf{v}}, \mathbf{x})$ with $\hat{\mathbf{v}}(y_0) = \psi(y_0, x_0, \hat{\mathbf{u}}(x_0)) = \psi(y_0, x_0, \mathbf{u}(x_0) - t_0)$. As t_0 ranges through $\mathbb{R}$, so does $\psi(y_0, x_0, \mathbf{u}(x_0) - t_0)$, giving the result.

B.4. *Proof of Lemma 1*

First, we prove the continuity of $\Psi : \mathbf{B}(X) \to \mathbf{B}(Y)$. The argument for the continuity of $\Phi : \mathbf{B}(Y) \to \mathbf{B}(X)$ is analogous.

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

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

Let (the following expressions are well-defined because $\mathbf{u}$ is bounded) $\bar{z} = \sup_{x\in X} \mathbf{u}(x) + 1$, $\underline{z} = \inf_{x\in X} \mathbf{u}(x) - 1$, and $Z = [\underline{z}, \bar{z}] \subset \mathbb{R}$. For every $\delta \in (0, 1)$ and $x \in X$, we then have

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

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

$$\|\tilde{\mathbf{u}} - \mathbf{u}\| < \delta \quad \Longrightarrow \quad \big|\psi\big(y, x, \tilde{\mathbf{u}}(x)\big) - \psi\big(y, x, \mathbf{u}(x)\big)\big| < \varepsilon'$$

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

$$\begin{aligned} &\big|\psi\big(y, x, \tilde{\mathbf{u}}(x)\big) - \psi\big(y, x, \mathbf{u}(x)\big)\big| < \varepsilon' \quad \text{for all } x \in X \text{ and } y \in Y \\ &\Longrightarrow \quad \sup_{y\in Y}\Big|\sup_{x\in X} \psi\big(y, x, \tilde{\mathbf{u}}(x)\big) - \sup_{x\in X} \psi\big(y, x, \mathbf{u}(x)\big)\Big| \le \varepsilon' < \varepsilon, \end{aligned}$$

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

Second, let $\mathcal{V} \subset \mathbf{B}(Y)$ be bounded, ensuring the existence of a compact interval $Z \subset \mathbb{R}$ such that $\mathbf{v}(Y) \subset Z$ holds for all $\mathbf{v} \in \mathcal{V}$. We then have $\Phi\mathbf{v}(x) \in [\min_{(x,y,v)\in X\times Y\times Z} \phi(x, y, v), \max_{(x,y,v)\in X\times Y\times Z} \phi(x, y, v)]$ for all $x \in X$ and $\mathbf{v} \in \mathcal{V}$, ensuring that $\Phi\mathcal{V} \subset \mathbf{B}(X)$ is bounded. The argument for Ψ is analogous.

B.5. *Proof of Corollary 1 and Completion of the Proof of Proposition 1*

We first use the defining property of a Galois connection (10) to establish (11)–(13) in the statement of Corollary 1.[1] In each case, we prove one of the two statements; the other statement follows by an analogous argument. First, for any $\mathbf{v} \in \mathbf{B}(Y)$ we trivially have $\Phi\mathbf{v} \geq \Phi\mathbf{v}$, so that setting $\mathbf{u} = \Phi\mathbf{v}$ in (10) yields (11). Second, let $\mathbf{v}_1 \geq \mathbf{v}_2$. By (11), we have $\mathbf{v}_2 \geq \Psi\Phi\mathbf{v}_2$ and thus $\mathbf{v}_1 \geq \Psi\Phi\mathbf{v}_2$. Applying (10) with $\mathbf{v} = \mathbf{v}_1$ and $\mathbf{u} = \Phi\mathbf{v}_2$ then gives the consequent of (12). Third, (11) gives $\mathbf{v} \geq \Psi\Phi\mathbf{v}$. Applying (12) with $\mathbf{v}_1 = \mathbf{v}$ and $\mathbf{v}_2 = \Psi\Phi\mathbf{v}$ to this inequality yields $\Phi\Psi\Phi\mathbf{v} \geq \Phi\mathbf{v}$. To establish the reverse inequality and hence (13), notice that, for every $\mathbf{v} \in \mathbf{B}(Y)$, we have $\Psi\Phi\mathbf{v} \geq \Psi\Phi\mathbf{v}$, so that using $\Psi\Phi\mathbf{v}$ in place of $\mathbf{v}$ and $\Phi\mathbf{v}$ in place of $\mathbf{u}$ in (10) yields the reverse inequality $\Phi\mathbf{v} \geq \Phi\Psi\Phi\mathbf{v}$.

We next show that (10) implies that Φ and Ψ are dualities that are dual to each other.

To confirm that Φ is a duality (with Ψ analogous), let $\underline{\mathbf{v}}$ be the infimum of some set $\mathcal{V} \subset \mathbf{B}(Y)$. Corollary 1(2) implies that $\Phi\underline{\mathbf{v}}$ then is an upper bound of $\Phi\mathcal{V}$. Let $\bar{\mathbf{u}}$ be any upper bound of $\Phi\mathcal{V}$. By (10), we then have $\mathbf{v} \geq \Psi\bar{\mathbf{u}}$ for all $\mathbf{v} \in \mathcal{V}$, implying $\underline{\mathbf{v}} \geq \Psi\bar{\mathbf{u}}$. Applying (10) again, this yields $\bar{\mathbf{u}} \geq \Phi\underline{\mathbf{v}}$, showing that $\Phi\underline{\mathbf{v}}$ is the supremum of $\Phi\mathcal{V}$. To see that Φ and Ψ are dual, note that (10) implies $\{\mathbf{u}|\mathbf{v} \geq \Psi\mathbf{u}\} = \{\mathbf{u}|\mathbf{u} \geq \Phi\mathbf{v}\}$, so that $\inf\{\mathbf{u}|\mathbf{v} \geq \Psi\mathbf{u}\} = \inf\{\mathbf{u}|\mathbf{u} \geq \Phi\mathbf{v}\} = \Phi\mathbf{v}$. An analogous argument establishes $\Psi\mathbf{u} = \inf\{\mathbf{v}|\mathbf{u} \geq \Phi\mathbf{v}\}$.

Finally, we argue that dualities that are dual to one another constitute a Galois connection. The proof is straightforward (cf. Singer (1997, p. 179)): Let $\mathbf{u} \geq \Phi\mathbf{v}$. Then $\Psi\mathbf{u} \leq \Psi\Phi\mathbf{v} \leq \inf\{\tilde{\mathbf{v}}|\Phi\tilde{\mathbf{v}} \leq \Phi\mathbf{v}\} \leq \mathbf{v}$, where the first inequality follows from the order-reversing property of the duality Ψ, the second inequality 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 (10); the other is analogous.

B.6. *Proof of Corollary 4*

We prove Corollary 4(1); Corollary 4(2) is analogous.

If $(\mathbf{u}, \mathbf{y})$ is implementable, there exists $\tilde{\mathbf{v}} \in \mathbf{B}(Y)$ implementing it, thus satisfying $\mathbf{u} = \Phi\tilde{\mathbf{v}}$, from which we obtain $\Psi\mathbf{u} = \Psi\Phi\tilde{\mathbf{v}}$. From the first inequality in (11) in Corollary 1(1), we have $\tilde{\mathbf{v}} \geq \Psi\Phi\tilde{\mathbf{v}}$ and thus $\tilde{\mathbf{v}} \geq \Psi\mathbf{u}$. Now suppose that $\Psi\mathbf{u}$ does not implement $\mathbf{y}$. Because $\Psi\mathbf{u}$ implements $\mathbf{u}$ (Corollary 3(1)), there exists $(\hat{x}, \hat{y}) \in X \times Y$ such that

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

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

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

resulting in a contradiction which finishes the proof.

[1] As noted in Birkhoff (1995, Section 5.8), the properties stated in (11)–(12) are in fact equivalent to (10) 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).

B.7. *Details for Remark 6*

We prove

$$\mathbf{v} \in \mathbf{I}(Y) \quad \Longleftrightarrow \quad Y_{\mathbf{v}} \text{ is nonempty-valued and onto;} \tag{B.4}$$

the proof of the other equivalence is analogous.

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

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

B.8. *Proof of Corollary 5*

We prove statements (1)–(3), with the proofs of the corresponding statements for $\mathbf{I}(Y)$ being analogous.

(1) Consider a sequence $(\mathbf{u}_n)_{n=1}^{\infty}$ of profiles in $\mathbf{I}(X)$ converging to some $\mathbf{u}^* \in \mathbf{B}(X)$. We want to show that $\mathbf{u}^*$ is implementable. For all $n \in \mathbb{N}$, let $\mathbf{v}_n = \Psi\mathbf{u}_n$. Because Ψ is continuous (Lemma 1), the sequence $(\mathbf{v}_n)_{n=1}^{\infty}$ converges to $\mathbf{v}^* = \Psi\mathbf{u}^*$. Corollary 3(1) implies that $\mathbf{v}_n$ implements $\mathbf{u}_n$, so that we have $\mathbf{u}_n = \Phi\mathbf{v}_n$ for all $n \in N$. Taking limits on both sides of this equation and using the continuity of Φ (Lemma 1), we obtain $\mathbf{u}^* = \Phi\mathbf{v}^*$. From Proposition 2, this establishes the implementability of $\mathbf{u}^*$, and hence that $\mathbf{I}(X)$ is closed. Next, suppose that the sequence $(\mathbf{u}_n)_{n=1}^{\infty}$ is in $\mathcal{U}_{\mathbf{y}} \subset \mathbf{I}(X)$. With the same construction of the sequence $(\mathbf{v}_n)_{n=1}^{\infty}$ as above, Corollary 4(1) then implies that $\mathbf{v}_n$ implements $\mathbf{y}$ for all n, so that

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

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

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

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

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

$$\|\hat{x} - x\| < \delta \quad \Longrightarrow \quad \|\mathbf{u}(\hat{x}) - \mathbf{u}(x)\| < \varepsilon \tag{B.5}$$

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

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 $\mathbf{v} \in \mathcal{V}$ implies $\underline{v} \leq \mathbf{v}(y) \leq \bar{v}$ 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

$$\|\hat{x} - x\| < \delta \quad \Longrightarrow \quad \|\phi(\hat{x}, y, v) - \phi(x, y, v)\| < \varepsilon \tag{B.6}$$

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

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

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

implying

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

where the outer inequalities are from (B.6) and the fact that $\underline{v} \leq \mathbf{v}(y) \leq \bar{v}$ holds for all $y \in Y$. Consequently, we have $\|\mathbf{u}(\hat{x}) - \mathbf{u}(x)\| < \varepsilon$, thus establishing (B.5).

(3) This follows from Corollary 5(2) and an application of the Arzela–Ascoli theorem (Ok (2007, p. 264)).

B.9. *Proof of Lemma 3*

We prove the first statement in the lemma; the second is analogous.

Fix an implementable $\mathbf{y} \in Y^X$ and consider $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}_\mathbf{y}$. Let $\mathbf{v}_1$ implement $(\mathbf{u}_1, \mathbf{y})$ and $\mathbf{v}_2$ implement $(\mathbf{u}_2, \mathbf{y})$. For any $x \in X$, we then have

$$\mathbf{u}_1(x) = \phi\big(x, \mathbf{y}(x), \mathbf{v}_1\big(\mathbf{y}(x)\big)\big), \tag{B.7}$$

$$\mathbf{u}_2(x) = \phi\big(x, \mathbf{y}(x), \mathbf{v}_2\big(\mathbf{y}(x)\big)\big). \tag{B.8}$$

From (B.7) and (B.8), it is immediate that

$$\mathbf{u}_1(x) \vee \mathbf{u}_2(x) = \phi\big(x, \mathbf{y}(x), \mathbf{v}_1\big(\mathbf{y}(x)\big) \wedge \mathbf{v}_2\big(\mathbf{y}(x)\big)\big) \tag{B.9}$$

holds for all $x \in X$. Combined with the equality $\Phi(\mathbf{v}_1 \wedge \mathbf{v}_2) = \mathbf{u}_1 \vee \mathbf{u}_2$ (cf. the first paragraph of Section 3.4.2), (B.9) shows that $\mathbf{v}_1 \wedge \mathbf{v}_2$ implements $(\mathbf{u}_1 \vee \mathbf{u}_2, \mathbf{y})$. Hence, $\mathbf{u}_1 \vee \mathbf{u}_2 \in \mathcal{U}_\mathbf{y}$.

From (B.7) and (B.8), it is also immediate that

$$\mathbf{u}_1(x) \wedge \mathbf{u}_2(x) = \phi\big(x, \mathbf{y}(x), \mathbf{v}_1\big(\mathbf{y}(x)\big) \vee \mathbf{v}_2\big(\mathbf{y}(x)\big)\big) \tag{B.10}$$

holds for all $x \in X$. From the implementation condition (4), we further have $\phi(x, y, \mathbf{v}_1(y)) \leq \mathbf{u}_1(x)$ and $\phi(x, y, \mathbf{v}_2(y)) \leq \mathbf{u}_2(x)$ for all $(x, y) \in X \times Y$, so that $\mathbf{u}_1(x) \wedge \mathbf{u}_2(x) \geq \phi(x, y, \mathbf{v}_1(y) \vee \mathbf{v}_2(y))$ holds for all x and y. Combined with (B.10), this shows that $\mathbf{v}_1 \vee \mathbf{v}_2$ implements $(\mathbf{u}_1 \wedge \mathbf{u}_2, \mathbf{y})$. Hence, $\mathbf{u}_1 \wedge \mathbf{u}_2 \in \mathcal{U}_\mathbf{y}$.

B.10. *Details for the Finite-Support Matching Models in Section 4.1.2*

With every finite-support matching model $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ satisfying Assumption 1 we associate a *matching model with a finite number of agents* as follows: there are finite sets of buyers $I = \{1, \dots, m\}$ and sellers $J = \{1, \dots, n\}$. Buyer i has type $x_i \in X$ and seller j has type $y_j \in Y$. Reservation utilities are given by $\underline{u}_i = \underline{\mathbf{u}}(x_i)$ for buyer $i \in I$ and $\underline{v}_j = \underline{\mathbf{v}}(y_j)$ for seller $j \in J$. The utility frontier available to a pair of matched agents $(i, j) \in I \times J$ is given by $\phi(x_i, y_j, v)$.

The standard definition of a match for such a matching model with a finite number of agents (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 $(i, j) \in I \times J$,

$\sum_{j\in J}\rho(i,j)\le 1$ for all $i\in I$, and $\sum_{i=I}\rho(i,j)\le 1$ for all $j\in J$. A stable outcome then consists of such a match and a specification of utility profiles $(u_1,\dots,u_n)$ and $(v_1,\dots,v_n)$ satisfying the natural counterparts to our feasibility and stability conditions (e.g., (20) becomes $u_i=\phi(x_i,y_j,v_j)$ for all (i,j) satisfying $\rho(i,j)=1$ and (25) becomes $u_i\ge\phi(x_i,y_j,v_j)$ for all $(i,j)\in I\times J$).

Every stable outcome for a matching model with a finite number of agents satisfies the equal treatment property (i.e., $x_i=x_{i'}$ implies $u_i=u_{i'}$ and $y_j=y_{j'}$ implies $v_j=v_{j'}$) if the characteristic function describing the utility frontier available to a pair of matched agents satisfies our Assumption 1. This allows us to identify stable outcomes for the matching model with a finite number of agents with stable outcomes for our finite-support matching model. Specifically, let $\mathcal{X}=\{x\in X|x=x_i \text{ for some } i\in I\}$ and $\mathcal{Y}=\{y\in Y|y=y_j \text{ for some } j\in J\}$ denote the supports of the type distributions in the finite-support matching model. For $x\in\mathcal{X}$, let $I(x)=\{i\in I|x_i=x\}$, and for $y\in\mathcal{Y}$, let $J(y)=\{j\in J|y_j=y\}$. Consider now a stable outcome $(\rho,u_1,\dots,u_m,v_1,\dots,v_n)$ for the matching model with a finite number of agents. Let $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ be arbitrary profiles in $\mathbf{B}(X)$ and $\mathbf{B}(Y)$. Given that equal treatment holds, setting

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

and

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

gives two well-defined profiles $\mathbf{u}\in\mathbf{B}(X)$ and $\mathbf{v}\in\mathbf{B}(Y)$. Let the measure λ have support contained in $\mathcal{X}\times\mathcal{Y}$ and on this set be 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 $(\lambda,\mathbf{u},\mathbf{v})$ is a stable outcome for the finite-support matching model.

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)). Hence, we may conclude that every finite-support matching model satisfying Assumption 1 has a stable outcome.

B.11. *Proof of Proposition 5(3)*

Let $(\lambda,\mathbf{u},\mathbf{v})$ be a pairwise stable outcome for the balanced matching model $(X,Y,\phi,\mu,\nu,\underline{\mathbf{u}},\underline{\mathbf{v}})$. Let $\mathcal{X}\subseteq X$ and $\mathcal{Y}\subseteq Y$ be the supports of μ and ν. Noticing that $\text{supp}(\lambda)\subseteq\mathcal{X}\times\mathcal{Y}$ holds, every pair of profiles $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ that satisfy $\tilde{\mathbf{u}}=\mathbf{u}$ on $\mathcal{X}$ and $\tilde{\mathbf{v}}=\mathbf{v}$ on $\mathcal{Y}$ satisfy (20) and (25), implying that, for any such pair, $(\lambda,\tilde{\mathbf{u}},\tilde{\mathbf{v}})$ is a pairwise stable outcome. It thus suffices to construct a pair of profiles satisfying $\tilde{\mathbf{u}}=\mathbf{u}$ on $\mathcal{X}$ and $\tilde{\mathbf{v}}=\mathbf{v}$ 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 $(x,y)\in\text{supp}(\lambda)$. (Otherwise, we would have $\lambda_X(\tilde{X})=0$ for some neighborhood $\tilde{X}$ of x, a contradiction.)

By (20) and (25), this implies that the restriction of the profile $\mathbf{v}$ to $\mathcal{Y}$ implements the restriction of the profile $\mathbf{u}$ to $\mathcal{X}$, that is,

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

Similarly, for every $y \in \mathcal{Y}$, there must exist $x \in \mathcal{X}$ with $(x, y) \in \operatorname{supp}(\lambda)$, so that (20) and (25) imply that restriction of $\mathbf{u}$ to $\mathcal{X}$ implements the restriction of $\mathbf{v}$ to $\mathcal{Y}$:

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

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

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

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

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

with sufficiently large $\breve{v}$ implements $\tilde{\mathbf{u}}$. Now, let $\tilde{\mathbf{v}} = \Psi\tilde{\mathbf{u}}$. As $\tilde{\mathbf{u}}$ is implementable, we then have that $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ implement each other (Corollary 3(1)). It remains to show that $\tilde{\mathbf{v}} = \mathbf{v}$ holds on $\mathcal{Y}$. This follows upon noting that (i) $\tilde{\mathbf{u}} = \mathbf{u}$ on $\mathcal{X}$ implies $\tilde{\mathbf{v}} \geq \mathbf{v}$ on $\mathcal{Y}$ (because the restriction of $\mathbf{u}$ to $\mathcal{X}$ implements the restriction of $\mathbf{v}$ to $\mathcal{Y}$) and (ii) we have $\tilde{\mathbf{v}} = \Psi\Phi\hat{\mathbf{v}}$, which implies (from Corollary 1(1)) $\hat{\mathbf{v}} \geq \tilde{\mathbf{v}}$ and therefore, because $\hat{\mathbf{v}} = \mathbf{v}$ on $\mathcal{Y}$, also implies the inequality $\mathbf{v} \geq \tilde{\mathbf{v}}$ on $\mathcal{Y}$.

B.12. *Proof of Lemma 5*

Suppose λ is a deterministic match satisfying $\lambda = \lambda_{\mathbf{y}}$ for an implementable $\mathbf{y}$. From Proposition 4(1), the implementability of $\mathbf{y}$ implies that there exists $\mathbf{u}$ and $\mathbf{v}$ implementing each other such that the graph of $\mathbf{y}$ is contained in $\Gamma_{\mathbf{u},\mathbf{v}}$. As the argmax correspondence $Y_{\mathbf{v}}$ is upper hemicontinuous (Corollary 2), its graph is closed. Hence, $\Gamma_{\mathbf{u},\mathbf{v}}$, which coincides with the graph of $Y_{\mathbf{v}}$ (Lemma 2), also contains the closure of the graph of $\mathbf{y}$. Moreover, the closure of the graph of $\mathbf{y}$ contains the support of $\lambda_{\mathbf{y}}$ (otherwise, there is a point (x, y) with a neighborhood that does not intersect the graph of $\mathbf{y}$ and which receives positive measure under $\lambda_{\mathbf{y}}$, a contradiction to the definition of $\lambda_{\mathbf{y}}$ in (28)). We thus have $\operatorname{supp}(\lambda) \subseteq \Gamma_{\mathbf{u},\mathbf{v}}$, implying that λ is pairwise stable (Propositions 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 $(\mathbf{u}, \mathbf{v})$ implementing each other such that $\operatorname{supp}(\lambda) \subseteq \Gamma_{\mathbf{u},\mathbf{v}}$. By Proposition 4(1), it remains to show that there exists a measurable assignment $\mathbf{y}$ with graph contained in $\Gamma_{\mathbf{u},\mathbf{v}}$ satisfying $\lambda_{\mathbf{y}} = \lambda$. By definition of a deterministic match, there exists a measurable assignment $\mathbf{y}'$ such that $\lambda = \lambda_{\mathbf{y}'}$ holds. If the graph of $\mathbf{y}'$ is contained in the support of λ, then we are done upon setting $\mathbf{y} = \mathbf{y}'$. It remains to consider the case that the graph of $\mathbf{y}'$ is not contained in the support of λ.

We construct the assignment $\mathbf{y}$. Let $\mathcal{X}$ denote the support of μ. First, we note that $\lambda_{\mathbf{y}'}$ does not depend on the specification of $\mathbf{y}'$ outside the support of μ. In addition, we can define the assignment $\mathbf{y}$ on $X \setminus \mathcal{X}$ so that $(x, \mathbf{y}(x)) \in \Gamma_{\mathbf{u},\mathbf{v}}$ holds for all $x \in X \setminus \mathcal{X}$.[2] Now let $\tilde{X} = \{x \in \mathcal{X} | (x, \mathbf{y}'(x)) \notin \operatorname{supp}(\lambda)\}$. The set $\tilde{X}$ is negligible (i.e., contained in a subset of $\mathcal{X}$ with measure zero) by definition of $\lambda_{\mathbf{y}'}$. Hence, we can complete the specification of $\mathbf{y}$ by taking $\mathbf{y}$ to equal a measurable selection from $\mathbf{Y}_{\mathbf{v}}$ (cf. footnote 2) (and hence $(x, \mathbf{y}(x)) \in \Gamma_{\mathbf{u},\mathbf{v}}$) on a subset of $\mathcal{X}$ that contains $\tilde{X}$ and has measure zero, and taking $\mathbf{y}$ to equal $\mathbf{y}'$ (and hence $(x, \mathbf{y}(x)) \in \operatorname{supp}(\lambda) \subseteq \Gamma_{\mathbf{u},\mathbf{v}}$) on the remainder of $\mathcal{X}$. This construction ensures that the graph of $\mathbf{y}$ is contained in $\Gamma_{\mathbf{u},\mathbf{v}}$. It follows immediately from the definitions of $\lambda_{\mathbf{y}}$ and $\lambda_{\mathbf{y}'}$ that we further have $\lambda_{\mathbf{y}} = \lambda_{\mathbf{y}'}$. As $\lambda_{\mathbf{y}'} = \lambda$ holds by assumption, this implies $\lambda_{\mathbf{y}} = \lambda$, finishing the proof.

B.13. *Proof of Corollary 6*

Fix a matching model $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ satisfying Assumption 1. We construct an augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$ 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 $\underline{\mathbf{u}}_0$ and $\underline{\mathbf{v}}_0$ duplicate $\underline{\mathbf{u}}$ on X and $\underline{\mathbf{v}}$ on Y, with $\underline{\mathbf{u}}(x_0) = \underline{\mathbf{v}}(y_0) = 0$.

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

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

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

Fourth, we let the measure μ_0 duplicate μ on the set X, and attach 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 $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$ is balanced.

The augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$ 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 $\underline{\mathbf{u}}$ and $\underline{\mathbf{v}}$ used in the construction of the extension of ϕ are (by assumption) continuous.

With any full match λ_0 for $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$, we associate the match λ for $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ obtained by restricting λ_0 to $X \times Y$. Vice versa, we can extend any match λ for $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ to a full match λ_0 for $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$ by assigning the masses of unmatched agents to the dummy agents and matching the remaining

[2]As we have noted earlier in this proof, the correspondence $\mathbf{Y}_{\mathbf{v}}$ 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)) $\tilde{\mathbf{y}}$. Take $\mathbf{y}$ to equal $\tilde{\mathbf{y}}$ on $X \setminus \mathcal{X}$.

masses of the dummy agents with each other. That is, we associate with λ the uniquely defined measure λ_0 satisfying

$$\lambda_0(\tilde{X} \times \{y_0\}) = \mu(\tilde{X}) - \lambda_X(\tilde{X}),$$
$$\lambda_0(\{x_0\} \times \tilde{Y}) = \nu(\tilde{Y}) - \lambda_Y(\tilde{Y}),$$

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

Because the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$ is balanced, we can invoke Proposition 6 to conclude that it has a pairwise stable outcome $(\lambda_0, \mathbf{u}_0, \mathbf{v}_0)$ satisfying $\mathbf{u}_0(x_0) = 0$. 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 B.2: *Let the matching model* $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ *satisfy Assumption* 1. *Then* $(\lambda, \mathbf{u}, \mathbf{v})$ *is a stable outcome of* $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ *if and only if it is associated with a pairwise stable outcome* $(\lambda_0, \mathbf{u}_0, \mathbf{v}_0)$, *satisfying* $\mathbf{u}_0(x_0) = 0$, *of the augmented matching model* $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$.

PROOF OF LEMMA B.2: Suppose the outcome $(\lambda_0, \mathbf{u}_0, \mathbf{v}_0)$ is a pairwise stable outcome of the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$ with $\mathbf{u}_0(x_0) = 0$ and let $(\boldsymbol{\lambda}, \mathbf{u}, \mathbf{v})$ be the associated outcome of $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$. 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 the full match λ_0 in the augmented matching model. Together with the equality $\mathbf{u}_0(x_0) = 0$, the feasibility condition (20) for types (x_0, y_0) in the augmented matching model then implies $\mathbf{v}_0(y_0) = 0$. For any type $x \in \text{supp}(\mu)$, (25) in the augmented matching model then implies $\mathbf{u}(x) \geq \phi_0(x, y_0, 0) = \underline{\mathbf{u}}(x)$ and similarly $\mathbf{v}(y) \geq \psi_0(y, x_0, 0) = \underline{\mathbf{v}}(y)$ for all $y \in \text{supp}(\nu)$. Thus, the participation constraints (23)–(24) in the associated outcome $(\lambda, \mathbf{u}, \mathbf{y})$ for the matching model hold. Next, the incentive constraints (25) in the augmented matching model,

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

imply

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

which are the incentive constraints in the matching model. It remains to check the feasibility conditions (20)–(22) to infer that $(\lambda, \mathbf{u}, \mathbf{v})$ is a stable outcome of $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$. As λ and λ_0 coincide on $X \times Y$, the feasibility conditions for the augmented matching model immediately imply $\mathbf{u}(x) = \phi(x, y, \mathbf{v}(y))$ for all (x, y) in the support of λ, which is (20). We then need only 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 $\mathbf{u}_0(x_0) = \mathbf{v}_0(y_0) = 0$ and the definition of ϕ_0) the feasibility condition

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

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

Conversely, suppose the outcome $(\lambda, \mathbf{u}, \mathbf{v})$ is a stable outcome of the matching model $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$. Let the profiles $\mathbf{u}_0 \in \mathbf{B}(X_0)$ and $\mathbf{v}_0 \in \mathbf{B}(Y_0)$ agree with $\mathbf{u}$ and $\mathbf{v}$ on X and Y and satisfy $\mathbf{u}_0(x_0) = 0$ and $\mathbf{v}_0(x_0) = 0$. Let λ_0 be the augmented match associated with λ. It suffices to show that $(\lambda_0, \mathbf{u}_0, \mathbf{v}_0)$ is a pairwise stable outcome of the matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$. The equalities $\mathbf{u}_0(x_0) = 0$ and $\mathbf{v}_0(y_0) = 0$ hold by construction. Feasibility and the conditions for pairwise stability follow from the feasibility and stability conditions for $(\lambda, \mathbf{u}, \mathbf{v})$ in the matching model $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ via arguments analogous to those establishing the previous direction. *Q.E.D.*

This completes the proof of Corollary 6.

B.14. *Proof of Proposition 8*

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

From Lemma B.2 in the proof of Corollary 6 in Appendix B.13, an outcome $(\lambda, \mathbf{u}, \mathbf{v})$ is stable in the matching model $(X, Y, \phi, \mu, \nu, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ if and only if the associated full outcome $(\lambda_0, \mathbf{u}_0, \mathbf{v}_0)$ is a pairwise stable outcome satisfying the initial condition $\mathbf{u}_0(x_0) = 0$ in the augmented matching model $(X_0, Y_0, \phi_0, \mu_0, \nu_0, \underline{\mathbf{u}}_0, \underline{\mathbf{v}}_0)$. Denote the set of pairwise stable buyer profiles satisfying the initial condition $\mathbf{u}_0(x_0) = 0$ in the augmented matching model by $\mathbb{U}_a$. With the obvious notational convention for the profile $(\mathbf{u}_0(x_0), \mathbf{u})$ of the augmented matching model, we then have $(\mathbf{u}_0(x_0), \mathbf{u}) \in \mathbb{U}_a$ if and only if both $\mathbf{u}_0(x_0) = 0$ and $\mathbf{u} \in \mathbb{U}_s$ hold. It is then immediate that $\mathbb{U}_s$ is a complete sublattice of $\mathbf{B}(X)$ if $\mathbb{U}_a$ is a complete sublattice of $\mathbf{B}(X_0)$.

To show that $\mathbb{U}_a$, which is nonempty by Proposition 6, is a complete sublattice of $\mathbf{B}(X_0)$, we first observe that $\mathbb{U}_a$ is the intersection of two closed sublattices of $\mathbf{B}(X_0)$, 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 $\mathbf{u}_0 \in \mathbf{B}(X_0)$ satisfying $\mathbf{u}_0(x_0) = 0$ (which is obviously a sublattice and closed). Hence, $\mathbb{U}_a$ is a closed sublattice of $\mathbf{B}(X_0)$. Further, the closed sublattice $\mathbb{U}_a$ is bounded, with the profile $\underline{\mathbf{u}}_0$ providing a lower bound and the profile $\Phi\underline{\mathbf{v}}_0$ providing an upper bound. Hence (Corollary 5(3)), $\mathbb{U}_a$ is a compact sublattice and therefore (by the same argument as in the proof of Lemma 4, cf. footnote 27 in Appendix A.2) complete.

B.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 $(\lambda, \mathbf{u}, \mathbf{v}) \in \mathbb{M} \times \mathbf{B}(X) \times \mathbf{B}(Y)$ be any triple satisfying the constraints in the principal's maximization problem defined in Section 5.1. Consider the triple $(\lambda, \mathbf{u}, \Psi\mathbf{u})$. The tariff $\Psi\mathbf{u}$ is implementable and implements $\mathbf{u}$ (Corollary 3(1)) and, further, implements any selection from $\mathbf{Y}_\mathbf{v}$ (Corollary 4(1)), so that $\mathbf{Y}_\mathbf{v}(x) \subseteq \mathbf{Y}_{\Psi\mathbf{u}}(x)$ holds for all $x \in X$. Consequently, we have $\mathrm{supp}(\lambda) \subseteq \Gamma_{\mathbf{u},\mathbf{v}} \subseteq \Gamma_{\mathbf{u},\Psi\mathbf{u}}$, ensuring that the triple $(\lambda, \mathbf{u}, \Psi\mathbf{u})$ is feasible in the principal's problem. As we have noted in the text following equation (30), the feasibility of $(\lambda, \mathbf{u}, \Psi\mathbf{u})$ implies that it results in the same expected payoff as $(\lambda, \mathbf{u}, \mathbf{v})$.

Step 2: From Step 1, we can restrict attention to $(\lambda, \mathbf{u}, \mathbf{v}) \in \mathbb{M} \times \mathbf{B}(X) \times \mathbf{I}(Y)$ when considering the principal's problem. As $\mathbf{v} \in \mathbf{I}(Y)$ implements $\mathbf{u} \in \mathbf{B}(Y)$ if and only if $\mathbf{u} =$

$\Phi\mathbf{v}$, we can eliminate the first constraint from the principal's problem and substitute this equality in the remaining constraints. The resulting problem is

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

Because implementable profiles are continuous (Proposition 2), the objective function in this problem is well-defined for all $\mathbf{v} \in \mathbf{I}(Y)$ and $\lambda \in \mathbb{M}$. Using (i) the definition of $F(\mathbf{v}, \lambda)$ in (31), observing (ii) that the constraint $\operatorname{supp}(\lambda) \subseteq \Gamma_{\Phi\mathbf{v},\mathbf{v}}$ is equivalent to $\lambda \in G(\mathbf{v})$, where $G(\mathbf{v})$ is defined in (32), and using (iii) the order reversal property of the implementation maps (Corollary 1(2)) to transform the constraint $\Phi\mathbf{v} \geq \underline{\mathbf{u}}$ into $\mathbf{v} \leq \Psi\underline{\mathbf{u}}$, we may rewrite the above problem as

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

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

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

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

Step 4: For $\mathbf{v} \in \mathbf{I}(Y)$, the correspondence $G(\mathbf{v})$ defined in (32) is nonempty-valued and compact-valued and upper hemicontinuous. To show that $G(\mathbf{v})$ is nonempty-valued, let $\mathbf{y}$ be a measurable selection (cf. footnote 2 in Appendix B.12) from $\mathbf{Y}_\mathbf{v}$ and let $\lambda_\mathbf{y}$ be the associated deterministic measure (cf. (28)). As $\mathbf{v}$ and $\Phi\mathbf{v}$ implement each other, the same argument as in the first paragraph of the proof of Lemma 5 yields that the support of $\lambda_\mathbf{y}$ is contained in $\Gamma_{\Phi\mathbf{v},\mathbf{v}}$. Hence, $G(\mathbf{v})$ is nonempty-valued.

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

Step 5: Fix $\mathbf{v} \in \mathbf{I}(Y)$ and consider the problem appearing in (33):

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

We have shown in Step 3 that $F(\mathbf{v}, \lambda)$ is continuous, and in Step 4 that $G(\mathbf{v})$ is nonempty-valued and compact-valued. Therefore, Weierstrass's extreme value theorem ensures that this problem has a solution so that the function $\Pi : \mathbf{I}(Y) \to \mathbb{R}$ 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

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

and let λ^* be an element of $\arg\max_{\lambda\in G(\mathbf{v}^*)} F(\mathbf{v}^*, \lambda)$. Then it is immediate from (33) that (v^*, λ^*) solves the problem

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

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

B.16. *Proof of Lemma 7*

We first construct an auxiliary balanced finite-support matching model $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}})$ 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 $\underline{\mathbf{u}}$ and $\underline{\mathbf{v}}$ by reservation utility profiles

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

and

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

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 $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}})$ (cf. Appendix B.10 of this supplement). By construction of $\underset{\sim}{\mathbf{u}}$ and $\underset{\sim}{\mathbf{v}}$, the inequalities $\phi(x_i, y_j, \underline{u}) > \phi(x_i, y_0, v_0) > \underline{u}$ hold for all $i, j \in \{1, \ldots, n\}$. 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 B.10) that the finite-support matching model $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underline{\mathbf{u}}, \underline{\mathbf{v}})$ has a fully matched stable outcome $(\hat{\lambda}_n, \mathbf{u}_n, \mathbf{v}_n)$ satisfying the initial condition $\mathbf{v}(y_0) = v_0$. 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 $(\hat{\lambda}_n, \mathbf{u}_n, \mathbf{v}_n)$ is also pairwise stable for the finite-support matching model $(X, Y, \phi, n \cdot \mu_n, n \cdot \nu_n, \underline{\mathbf{u}}, \underline{\mathbf{v}})$. Letting $\lambda_n = \hat{\lambda}_n / n$, it is obvious that $(\lambda_n, \mathbf{u}_n, \mathbf{v}_n)$ is a pairwise stable outcome for the matching model $(X, Y, \phi, \mu_n, \nu_n, \underline{\mathbf{u}}, \underline{\mathbf{v}})$. Finally, from Proposition 5(3), we may assume that $\mathbf{u}_n$ and $\mathbf{v}_n$ implement each other, giving a pairwise stable outcome $(\lambda_n, \mathbf{u}_n, \mathbf{v}_n)$ satisfying all the conditions from the statement of the lemma.

APPENDIX C: EXAMPLES

C.1. *Example 1: The Set of Implementable Profiles Is not a Sublattice*

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

$$\phi(x, 1, v) = 1 - v,$$
$$\phi(x, 2, v) = 2 + x - v$$

for $x \in X$. The inverse generating function then is

$$\psi(1, x, u) = 1 - u,$$
$$\psi(2, x, u) = 2 + x - u.$$

The profiles $\mathbf{u}_1 = (1, 1, 1)$ and $\mathbf{u}_2 = (0, 1, 2)$ are both implementable ($\mathbf{v}_1 = (0, 4)$ implements $\mathbf{u}_1$ and $\mathbf{v}_2 = (1, 3)$ implements $\mathbf{u}_2$). The profile $\mathbf{u}_1 \wedge \mathbf{u}_2 = (0, 1, 1)$, however, is not implementable. Hence, $\mathbf{I}(X)$ is not a sublattice of $\mathbf{B}(X)$. To establish that $\mathbf{u}_1 \wedge \mathbf{u}_2 = (0, 1, 1)$ is not implementable, it suffices to note (Remark 6) that $\mathbf{X}_{(0,1,1)}$ is not onto: $x = 1$ is the unique maximizer of $\psi(1, x, \mathbf{u}(x))$ and $x = 3$ is the unique maximizer of $\psi(2, x, \mathbf{u}(x))$. (Alternatively, we may note that $\Psi(0, 1, 1) = (0, 4) = \mathbf{v}_1$. As $\mathbf{v}_1$ implements $\mathbf{u}_1 = (1, 1, 1)$, we obtain $\Phi\Psi(0, 1, 1) \neq (0, 1, 1)$ with Proposition 3(1) then implying that $(0, 1, 1)$ is not implementable.)

C.2. *Example 2: The Participation Constraint Is not Binding in a Solution to the Principal's Problem*

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

$$\phi(1, 1, v) = 3 - 2v,$$
$$\phi(1, 2, v) = 2 - v,$$

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

Let $\mu(1) = \mu(2) = 1/2$ and $\underline{\mathbf{u}}(1) = \underline{\mathbf{u}}(2) = 0$. 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 $\mathbf{y} = (\mathbf{y}(1), \mathbf{y}(2))$.

Figure S.1 illustrates the set of profiles $\mathbf{v} = (\mathbf{v}(1), \mathbf{v}(2))$ and, for each such profile, identifies the assignment(s) $\mathbf{y} = (\mathbf{y}(1), \mathbf{y}(2))$ 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 $\mathbf{B}(Y)$ of profiles $\mathbf{v}$. All assignments $\mathbf{y} \in Y^X$ are implementable, but only the constant assignments $\mathbf{y} = (1, 1)$ and $\mathbf{y} = (2, 2)$ are strongly implementable.

The set of implementable tariffs $\mathbf{I}(Y)$ is the (blue and orange, or dark and light) shaded area in Figure S.1, 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 $\mathbf{v}(2) \leq 2$ satisfy $\Phi\mathbf{v} \geq \underline{\mathbf{u}}$, whereas tariffs in the shaded area of Figure S.1 with $\mathbf{v}(2) > 2$ lead to a violation of agent 1's participation constraint. Hence, the set $\{\mathbf{v} \in \mathbf{I}(Y) : \mathbf{v} \geq \Psi\underline{\mathbf{u}}\}$ appearing in the nonlinear pricing problem (34) is given by that portion of the shaded area in Figure S.1 for which $\mathbf{v}(2) \leq 2$.

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 $\mathbf{y} = (2, 2)$ or $\mathbf{y} = (2, 1)$ by choosing $\mathbf{v} = (3, 2)$, she could implement $\mathbf{y} = (1, 1)$ by choosing $(1.5, 2)$, or she could implement $\mathbf{y} = (1, 2)$ by choosing $\mathbf{v} = (1, 1)$. Now, sup-

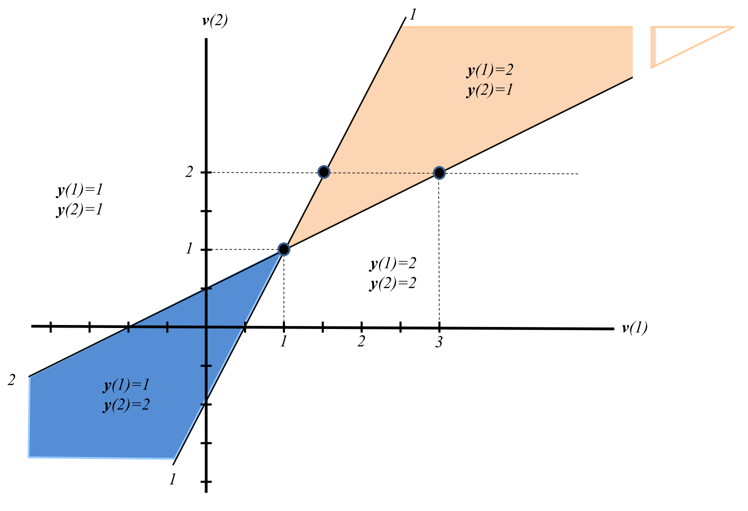

FIGURE S.1.—Illustration of the assignments $\mathbf{y}$ implemented by various profiles $\mathbf{v}$, the set $\mathbf{I}(Y)$ of implementable profiles (colored or shaded areas, including the boundary), and the feasible set for the principal's nonlinear pricing problem (the portion of the shaded areas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

pose the principal's utility function is

$$\pi(1, 1, v) = v + 5,$$
$$\pi(1, 2, v) = v,$$
$$\pi(2, 1, v) = v,$$
$$\pi(2, 2, v) = v + 5.$$

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

This example 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 $\mathbf{y} = (1, 2)$ to $\mathbf{y} = (2, 1)$ affects the principal's utility only through the change in tariff, ensuring that the principal would welcome the attendant increase in tariff from implementing $\mathbf{y} = (2, 1)$ with the tariff $\mathbf{v} = (3, 2)$ rather than implementing $\mathbf{y} = (1, 2)$ with the tariff $\mathbf{v} = (1, 1)$.

APPENDIX D: EXTENSIONS

D.1. *Exclusion in the Principal-Agent Model*

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 nonparticipation 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 not contained in Y, 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

$$\phi_0(x, y_0, 0) = \underline{\mathbf{u}}(x). \tag{D.1}$$

Hence, in the absence of a transfer ($v = 0$), agent types choosing the outside option y_0 receive their reservation utility $\underline{\mathbf{u}}(x)$. 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

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

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 $(X, Y_0, \phi_0, \mu, \pi_0, \underline{\mathbf{u}})$ 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 $(X, Y_0, \phi_0, \mu, \pi_0, \underline{\mathbf{u}})$, it is immediate from Proposition 9 that the principal-agent model with exclusion has a solution $(\lambda, \mathbf{u}, \mathbf{v})$ in which $\mathbf{u}$ and $\mathbf{v}$ implement each other. Further, because any such solution respects the participation constraint $\mathbf{u} \geq \underline{\mathbf{u}}$, it satisfies the constraint that the principal cannot charge the agent for choosing the outside option.[3]

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

Provided that the participation constraint binds for some type of agent in a solution to the principal-agent model with exclusion, we must have $\mathbf{v}(y_0) = 0$, 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 $\phi_0(x, y_0, v) = \underline{\mathbf{u}}(x) - v$), then the principal-agent model with exclusion will satisfy quasilinearity. As in Jullien's quasilinear model of exclusion, there is then no loss of generality to restrict the principal to tariffs satisfying $\mathbf{v}(y_0) = 0$ (Jullien (2000, footnote 7)).[4]

If the participation constraint does not hold with equality for any agent type in a solution to the principal-agent model with exclusion, then such a solution might satisfy $\mathbf{v}(y_0) < 0$. 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 (in Appendix C.2 of this supplement), all types of the agent obtain strictly higher utility than their reservation utility. In this case, the optimal $(\mathbf{u}, \mathbf{y})$ can also be implemented by a (non-implementable) tariff $\mathbf{v}$ satisfying $\mathbf{v}(y_0) = 0$. The second, more interesting, case is that some excluded type receives the strictly positive payment $-\mathbf{v}(y_0)$ 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 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

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

[3]Using the obvious notation for the inverse generating function and the implementation map in the model with exclusion, the formal argument is this: If $\mathbf{u}$ and $\mathbf{v}$ implement each other, the participation constraint implies $\mathbf{v} \leq \Psi_0 \underline{\mathbf{u}}$. Therefore, we have $\mathbf{v}(y_0) \leq \psi_0(y_0, x, \underline{\mathbf{u}}(x))$ for all $x \in X$. From (D.1), the right side of the latter inequality is equal to zero.

[4]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.

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

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

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 $\mathbf{v}(1) = 2/3 = -\mathbf{v}(0)$. Hence, the principal pays agent 2 to not participate.

D.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 than 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 $(X, Y, \phi, \mu, \nu, \pi, \underline{\mathbf{u}})$ 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:

$$\phi_\Delta(x, \zeta, v) = \int_Y \phi(x, y, v)\, d\zeta(y),$$
$$\pi_\Delta(x, \zeta, v) = \int_Y \pi(x, y, v)\, d\zeta(y),$$

thereby obtaining a principal-agent model $(X, \Delta Y, \phi_\Delta, \mu, \pi_\Delta, \underline{\mathbf{u}})$ 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.[5] In this model, our version of the taxation principle (Remark 1) as well as all the results from Section 5 continue to hold.

[5]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. Consider continuity. 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

$$\int_Y \phi(x, y, v)\, d\zeta_n(y) \quad \to \quad \int_Y \phi(x, y, v)\, d\zeta(y). \tag{D.2}$$

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 Z of $\mathbb{R}$. Then, for any ε, there exists a sufficiently large N such

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 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.[6] Therefore, the 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.

D.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 $\mathbf{v}(y)$ to the principal. A decision y now is a function $\mathbf{w} : Z \to [\underline{w}, \overline{w}]$ identifying, for each output level z, the wage $\mathbf{w}(z) \in [\underline{w}, \overline{w}]$ 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 $z - (w - v)$.

---

that, for all $n \geq N$,

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

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

[6]For example, let $q \in [0, \bar{q}]$ be the quantity of some good. Ordinarily, we would take $Y = [0, \bar{q}]$ and then suppose that a monopolistic seller (the principal) with utility function $\pi(x, q, v)$ designs a tariff specifying payments $\mathbf{v}(q)$ 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 $\mathbf{v}(q, t)$. 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 $(X, \hat{Y}, \hat{\phi}, \mu, \hat{\pi}, \underline{\mathbf{u}})$ that satisfies Assumptions 1 and 3 if the original model $(X, Y, \phi, \mu, \pi, \underline{\mathbf{u}})$ does so and describes the same underlying economic environment.

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

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

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

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

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 $\pi(x, \mathbf{w}, v)$ 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 $\mathbf{w}$. Our results will apply as long as attention is restricted to circumstances in which the set Y can reasonably be taken to be compact.

REFERENCES

ALIPRANTIS, C., AND K. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide* (Third Ed.). Springer-Verlag. [9]

BIRKHOFF, G. (1995): *Lattice Theory*. (Third Ed.) Colloquium Publications Vol. 25. American Mathematical Society. [4]

DEMANGE, G., AND D. GALE (1985): "The Strategy Structure of Two-Sided Matching Markets," *Econometrica*, 53 (4), 873–888. [14]

FERRERA, J. (2014): *An Introduction to Nonsmooth Analysis*. Elsevier. [1]

JULLIEN, B. (2000): "Participation Constraints in Adverse Selection Problems," *Journal of Economic Theory*, 93 (1), 47. [16,17]

KADAN, O., P. J. RENY, AND J. SWINKELS (2017): "Existence of Optimal Mechanisms in Principal-Agent Problems," *Econometrica*, 85 (3), 769–823. [19]

LAFFONT, J.-J., AND D. MARTIMORT (2002): *The Theory of Incentives*. Princeton University Press. [19]

LAFFONT, J.-J., AND J. TIROLE (1993): *A Theory of Incentives in Procurement and Regulation*. MIT Press. [19]

MYERSON, R. B. (1982): "Optimal Coordination in Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10 (1), 67–81. [19]

OK, E. A. (2007): *Real Analysis With Economic Applications*. Princeton University Press. [6,13]

ORE, O. (1944): "Galois Connexions," *Trans. Amer. Math. Soc*, 55 (3), 493–513. [4]

ROTH, A. E., AND M. A. O. SOTOMAYOR (1990): *Two-Sided Matching*. Cambridge University Press. [6,7]

SHIRYAEV, A. N. (1996): *Probability* (Second Ed.). Springer-Verlag. [13]

SINGER, I. (1997): *Abstract Convex Analysis*. Wiley-Interscience. [4]

STRAUSZ, R. (2006): "Deterministic versus Stochastic Mechanisms in Principal-Agent Models," *Journal of Economic Theory*, 128 (1), 306–314. [18]

Co-editor Joel Sobel handled this manuscript.

Manuscript received 16 March, 2015; final version accepted 2 April, 2018; available online 3 April, 2018.