Supplement to "A note on identification of discrete choice models for bundles and binary games"

(*Quantitative Economics*, Vol. 8, No. 3, November 2017, 1021–1036)

Jeremy T. Fox
Department of Economics, Rice University and NBER

Natalia Lazzati
Department of Economics, University of California, Santa Cruz

These appendixes contain three sets of results. In Appendix B, we include some theoretical tools we use to prove Lemma 1: Topkis's theorem and stochastic dominance. In Appendix C, we show by example the key role of exclusion restrictions in our analysis. In Appendix D, we extend our identification results to the case of three or more goods or players.

Appendix B: Monotone comparative statics

The proof of Lemma 1 relies on Topkis' theorem and the concept of stochastic dominance.

Topkis' theorem (Topkis (1998)). *Let* $f(a_1, a_2, x) : \mathcal{A}_1 \times \mathcal{A}_2 \times \mathbb{R} \to \mathbb{R}$, *where* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *are finite ordered sets. Assume that* $f(a_1, a_2, x)$ (i) *is supermodular in* (a_1, a_2) *and that* (ii) *it has increasing differences in* (a_1, x) *and* (a_2, x). *Then* $\arg\max\{f(a_1, a_2, x) \mid (a_1, a_2) \in \mathcal{A}_1 \times \mathcal{A}_2\}$ *increases in* x *with respect to the strong set order.*[1] (*According to this order, we write* $\mathcal{A} \geq_S \mathcal{B}$ *if for every* $a \in \mathcal{A}$ *and* $b \in \mathcal{B}$, *we have that* $a \vee b \in \mathcal{A}$ *and* $a \wedge b \in \mathcal{B}$.)

The concept of first order (or standard) stochastic dominance (FOSD) is based on upper sets. Let us consider $(\Omega, \geq)$, where Ω is a set and $\geq$ defines a partial order on it. A subset $U \subset \Omega$ is an upper set if $x \in U$ and $x' \geq x$ imply $x' \in U$.

Jeremy T. Fox: jeremyfox@gmail.com
Natalia Lazzati: nlazzati@ucsc.edu

[1] For any two elements $a, a' \in \mathcal{A}_1 \times \mathcal{A}_2$ we write $a \vee a'$ $(a \wedge a')$ for the least upper bound (greatest lower bound). We say $f(a_1, a_2, x)$ is supermodular in (a_1, a_2) if, for all $a, a' \in \mathcal{A}_1 \times \mathcal{A}_2$,

$$f(a \vee a', x) + f(a \wedge a', x) \geq f(a, x) + f(a', x).$$

We say that $f(a_1, a_2, x)$ has increasing differences in (a_1, x) if, for all $a_1' > a_1$ and $x' > x$,

$$f(a_1', a_2, x') - f(a_1, a_2, x') \geq f(a_1', a_2, x) - f(a_1, a_2, x).$$

DOI: 10.3982/QE489

First order stochastic dominance

Let $X', X \in \mathbb{R}^n$ be two random vectors. We say X' is higher than X with respect to first order stochastic dominance if, for all upper sets $U \subset \mathbb{R}^n$,

$$\Pr\big(X' \in U\big) \geq \Pr(X \in U).$$

Appendix C: Nonidentification without exclusion restrictions

We show using a simple example that we cannot recover the joint distribution of unobservables $F(\varepsilon_1, \varepsilon_2)$ without the exclusion restrictions, captured by the roles of Z_1 and Z_2 in

$$U(a, W, Z, \varepsilon, \eta) = \sum_{i=1,2} \big(u_i(W) + Z_i + \varepsilon_i\big) \cdot a_i + \eta \cdot v(W) \cdot a_1 \cdot a_2. \tag{A.1}$$

To this end, assume that $Z_1 = Z_2 = Z$, where Z is a scalar random variable. Thus Z is not an excluded variable.

We focus on a simpler version of our model. Specifically, we drop W and assume both no interaction effects and zero homogeneous stand-alone payoffs. That is, the researcher knows that $\eta \cdot v = 0$, $u_1 = 0$, and $u_2 = 0$. Because $\eta \cdot v = 0$, the distribution of η is not identified and indeed is not an interesting object to recover. If $F(\varepsilon_1, \varepsilon_2)$ is not identified under these strong conditions, then it is not identified if we allow for v, u_1, and u_2 to differ from zero. Under these conditions, the choice probabilities simplify to

$$\Pr\big((1, 1) \mid z\big) = \Pr(\varepsilon_1 \geq -z, \varepsilon_2 \geq -z \mid z),$$
$$\Pr\big((0, 0) \mid z\big) = \Pr(\varepsilon_1 \leq -z, \varepsilon_2 \leq -z \mid z),$$
$$\Pr\big((1, 0) \mid z\big) = \Pr(\varepsilon_1 \geq -z, \varepsilon_2 \leq -z \mid z),$$
$$\Pr\big((0, 1) \mid z\big) = \Pr(\varepsilon_1 \leq -z, \varepsilon_2 \geq -z \mid z).$$

In this case, choice probabilities in the data only reveal the probability that the unobservables $(\varepsilon_1, \varepsilon_2)$ belong to rectangles with the common vertex $(-z, -z)$. Note that if Z has large support, then the marginal distributions of ε_1 and ε_2 are immediately identified by treating the decisions to buy items 1 and 2 as separate binary choice problems (Manski (1988)). This approach is possible as we assumed $\eta \cdot v$ is 0. Because the marginal distributions are identified, the nonidentification of $F(\varepsilon_1, \varepsilon_2)$ must arise from the copula linking the two marginals to create the joint distribution $F(\varepsilon_1, \varepsilon_2)$—which exists by Sklar's theorem.

We show the nonidentification of $F(\varepsilon_1, \varepsilon_2)$ by describing *two different joint distributions*, with the same marginals, that give the *same choice probabilities*. Let $F^1(\varepsilon_1, \varepsilon_2)$ be the CDF of a uniform distribution on $[-1, 1] \times [-1, 1]$, which has a probability density function (p.d.f.) of 1/4 on $[-1, 1] \times [-1, 1]$. The choice probabilities for each $z \in [-1, 1]$ sum up to 1 and are given by

$$\Pr\big((1, 1) \mid z\big) = \Pr(\varepsilon_1 \geq -z, \varepsilon_2 \geq -z \mid z) = (1 + z) \cdot (1 + z) \cdot \frac{1}{4},$$

$$\Pr\big((0,0)\mid z\big)=\Pr(\varepsilon_1\le -z,\varepsilon_2\le -z\mid z)=(1-z)\cdot(1-z)\cdot\frac{1}{4},$$
$$\Pr\big((1,0)\mid z\big)=\Pr(\varepsilon_1\ge -z,\varepsilon_2\le -z\mid z)=(1+z)\cdot(1-z)\cdot\frac{1}{4},$$
$$\Pr\big((0,1)\mid z\big)=\Pr(\varepsilon_1\le -z,\varepsilon_2\ge -z\mid z)=(1-z)\cdot(1+z)\cdot\frac{1}{4}.$$

Let us next construct $F^2(\varepsilon_1,\varepsilon_2)$ from $F^1(\varepsilon_1,\varepsilon_2)$ by shifting mass uniformly from each half-quadrant triangle to the next one clockwise, so that, under $F^2(\varepsilon_1,\varepsilon_2)$,

$$\begin{aligned}
\Pr(\varepsilon_2\ge\varepsilon_1, 0\le\varepsilon_1\le 1, 0\le\varepsilon_2\le 1)&=0\quad\text{and}\\
\Pr(\varepsilon_2\le\varepsilon_1, 0\le\varepsilon_1\le 1, 0\le\varepsilon_2\le 1)&=1/4,\\
\Pr(-\varepsilon_2\le\varepsilon_1, 0\le\varepsilon_1\le 1, -1\le\varepsilon_2\le 0)&=0\quad\text{and}\\
\Pr(-\varepsilon_2\ge\varepsilon_1, 0\le\varepsilon_1\le 1, -1\le\varepsilon_2\le 0)&=1/4,\\
\Pr(-\varepsilon_1\le-\varepsilon_2, -1\le\varepsilon_1\le 0, -1\le\varepsilon_2\le 0)&=0\quad\text{and}\\
\Pr(-\varepsilon_1\ge-\varepsilon_2, -1\le\varepsilon_1\le 0, -1\le\varepsilon_2\le 0)&=1/4,\\
\Pr(-\varepsilon_1\ge\varepsilon_2, -1\le\varepsilon_1\le 0, 0\le\varepsilon_2\le 1)&=0\quad\text{and}\\
\Pr(-\varepsilon_1\le\varepsilon_2, -1\le\varepsilon_1\le 0, 0\le\varepsilon_2\le 1)&=1/4.
\end{aligned}$$

Because the shift is uniform, the p.d.f. of $(\varepsilon_1,\varepsilon_2)$ in the region of positive probability is $1/2$.

We next show that while $F^2(\varepsilon_1,\varepsilon_2)$ differs from $F^1(\varepsilon_1,\varepsilon_2)$, they generate the same choice probabilities in the data; a similar argument can be used to show that they share the same marginals. The choice probabilities utilize the formula for the area of a right triangle, $\frac{1}{2}$ times base times height. The choice probabilities under $F^2(\varepsilon_1,\varepsilon_2)$ if $z\le 0$ (or $-z\ge 0$) are

$$\begin{aligned}
\Pr\big((1,1)\mid z\big)&=\Pr(\varepsilon_1\ge -z,\varepsilon_2\ge -z\mid z)=\frac{(1+z)\cdot(1+z)}{2}\cdot\frac{1}{2}=(1+z)\cdot(1+z)\cdot\frac{1}{4},\\
\Pr\big((0,0)\mid z\big)&=\Pr(\varepsilon_1\le -z,\varepsilon_2\le -z\mid z)\\
&=\frac{1}{4}+\frac{(-z)\cdot(-z)}{2}\cdot\frac{1}{2}+(1+z)\cdot(-z)\cdot\frac{1}{2}+\frac{(-z)\cdot(-z)}{2}\cdot\frac{1}{2}+\frac{(-z)\cdot(-z)}{2}\cdot\frac{1}{2}\\
&=(1-z)\cdot(1-z)\cdot\frac{1}{4},\\
\Pr\big((1,0)\mid z\big)&=\Pr(\varepsilon_1\ge -z,\varepsilon_2\le -z\mid z)=(1+z)\cdot(-z)\cdot\frac{1}{2}+\frac{(1+z)\cdot(1+z)}{2}\cdot\frac{1}{2}\\
&=(1+z)\cdot(1-z)\cdot\frac{1}{4},\\
\Pr\big((0,1)\mid z\big)&=\Pr(\varepsilon_1\le -z,\varepsilon_2\ge -z\mid z)=(1+z)\cdot(-z)\cdot\frac{1}{2}+\frac{(1+z)\cdot(1+z)}{2}\cdot\frac{1}{2}\\
&=(1-z)\cdot(1+z)\cdot\frac{1}{4}.
\end{aligned}$$

The algebra for the case where $z \geq 0$ (or $-z \leq 0$) is similar, so we omit it. Since two different CDFs $F^1(\varepsilon_1, \varepsilon_2)$ and $F^2(\varepsilon_1, \varepsilon_2)$ generate the same choice probabilities for all z, the joint distribution of $\varepsilon_1, \varepsilon_2$ cannot be uniquely recovered from the data.

Appendix D: Three or more goods and players

D.1 *Identification results for a general model*

Consider an agent who faces $\mathcal{N} = \{1, 2, \dots, n\}$ binary choice variables that are not mutually exclusive. Thus, her choice set is $\{0, 1\}^n$. We define $a = (a_i)_{i \leq n}$.

If the agent selects only variable i, then her payoff is $u_i(W) + Z_i + \varepsilon_i$, where $W \in \mathbb{R}^k$ is a vector of explanatory variables and $\varepsilon = (\varepsilon_i)_{i \leq n} \in \mathbb{R}^n$ indicates a vector of random terms distributed according to $F_{\varepsilon|W,Z}$ that is observed by the agent but not by the econometrician. We define $Z = (Z_i)_{i \leq n}$, where Z_i is a scalar for $i = 1, 2, \dots, n$. As before, Z represents the excluded variables. We write $S(a) = \{i \in \mathcal{N} : a_i = 1\}$ for the set of variables that are actually selected if vector a is chosen. The agent selects the action profile a to maximize her utility,

$$U(a, W, Z, \varepsilon) = \sum_{i \leq n} \big(u_i(W) + Z_i + \varepsilon_i\big) \cdot a_i + v\big(S(a), W\big), \tag{A.2}$$

where $v(S(a), W)$ is the interaction effect between the selected variables. This specification allows the interaction term to vary with the items selected. We normalize $v(S(a), W)$ so that it is 0 when $|S(a)| < 2$ and let the overall utility be 0 if $a = (0, \dots, 0)$ is chosen.

Our purpose is to identify $((u_i)_{i \leq n}, v, F_{\varepsilon|W,Z})$ from available choice data $\Pr(a \mid w, z)$. We next provide the set of identifying restrictions. Let Z_{-i} be all excluded explanatory variables other than that for the binary variable i.

B1. The term $Z_i \mid W, Z_{-i}$ has support on all $\mathbb{R}$ for $i = 1, 2, \dots, n$.

B2. We have (i) $F_{\varepsilon|W,Z} = F_{\varepsilon|W}$ and (ii) $E(\varepsilon \mid W) = (0, 0, \dots, 0)$.

B3. The term $\varepsilon \mid W$ has an everywhere positive Lebesgue density on its support.

B4. For each $W = w$, there exists a known vector $\widehat{a}^w \in \{0, 1\}^n$ with the following property: For all z and ε, $U(a, w, z, \varepsilon)$ is maximized at $\widehat{a}^w$ if $U(\widehat{a}^w_i, \widehat{a}^w_{-i}, w, z, \varepsilon) \geq U(a_i, \widehat{a}^w_{-i}, w, z, \varepsilon)$ for all $i \in \mathcal{N}$, where $a_i = 1 - \widehat{a}^w_i$.

We do not allow for heterogeneity in the interaction terms in the multiple binary variables model. In addition, we need to add B4. This assumption requires the existence of a vector of choices (known by the econometrician) such that if the vector is a local maximizer, then it is a global maximizer as well. We provide sufficient conditions for B4 below.

Theorem D.1. *Under* B1–B4, $((u_i(\cdot))_{i \leq n}, v(\cdot), F_{\varepsilon|W})$ *is identified.*

The proof of Theorem D.1, in Appendix D.4 below, first uses condition B4 to trace $F_{\varepsilon|W}$ using variation in Z. We then use the known $F_{\varepsilon|W}$ to show—via contradiction—that there exist realizations of Z where different values of $((u_i)_{i \leq n}, v)$ lead to different bundle choice probabilities.

The next four conditions are sufficient for assumption B4 to hold.

Two binary variables. Assumption B4 always holds if there are two choice variables, so that B4 in this case is not stronger than the analysis in the main text. If B1–B3 are satisfied, then, by Lemma 1, the sign of $v(w) = v((1,2), w)$ is identified for each $W = w$. It can be easily shown that B4 holds with $\widehat{a}^w = (1,0)$ or $\widehat{a}^w = (0,1)$ when $v(w) \geq 0$ and $\widehat{a}^w = (0,0)$ or $\widehat{a}^w = (1,1)$ when $v(w) \leq 0$.

Negative interaction effects. The second sufficient condition relies on the items being substitutes. This holds if $U(a, w, z, \varepsilon)$ has the negative single-crossing property in $(a_i; a_{-i})$ for all $i \in \mathcal{N}$; that is, if, for all $a'_{-i} > a_{-i}$ (in the coordinatewise order) and w, z, ε, we have

$$\begin{aligned} &U(a_i = 1, a_{-i}, w, z, \varepsilon) - U(a_i = 0, a_{-i}, w, z, \varepsilon) \leq (<) 0 \quad \Longrightarrow \\ &U\big(a_i = 1, a'_{-i}, w, z, \varepsilon\big) - U\big(a_i = 0, a'_{-i}, w, z, \varepsilon\big) \leq (<) 0. \end{aligned}$$

In this case, B4 holds with $\widehat{a}^w = (0, 0, \ldots, 0)$ for each w.

Positive or mixed interaction effects. The third sufficient condition applies to the case where the binary variables are all complements or the interaction effects have different signs. Though this result extends to the case of an arbitrary (finite) number of binary choice variables, we present the sufficient condition for the case of three to simplify the exposition. In this case, B4 holds with $\widehat{a}^w = (1, 0, 0)$ for each $W = w$ if

$$\begin{aligned} v\big((1,2), w\big) &\geq 0, \\ v\big((1,3), w\big) &\geq 0, \\ v\big((1,2), w\big) + v\big((1,3), w\big) &\geq v\big((2,3), w\big), \\ v\big((1,2), w\big) + v\big((1,3), w\big) &\geq v\big((1,2,3), w\big). \end{aligned}$$

The first two conditions require that binary variable 1 be a complement with each of the other two variables. The third condition requires that the sum of the interaction effects among variable 1 and the other two be larger than the interaction effect between the latter; notice that this condition always holds if the choice variables 2 and 3 are mutual substitutes. The last condition is a subadditivity condition. Though we selected choice variable 1, the same idea applies if the leading choice variable is either 2 or 3.

Global concavity for discrete domains. Assumption B4 is a local notion of concavity for discrete domains. Thus, it also holds under a global analog of discrete concavity; that is, if, for all $a, a' \in \{0,1\}^n$ with $\|a - a'\| = 2$,

$$\max_{a'' : \|a - a''\| = \|a' - a''\| = 1} U\big(a'', w, z, \varepsilon\big) \begin{cases} > \min\big\{U(a, w, z, \varepsilon), U\big(a', w, z, \varepsilon\big)\big\}, \\ \quad \text{if } U(a, w, z, \varepsilon) \neq U\big(a', w, z, \varepsilon\big), \\ \geq U(a, w, z, \varepsilon) = U\big(a', w, z, \varepsilon\big), \\ \quad \text{otherwise.} \end{cases}$$

In this case, B4 holds with any $\widehat{a}^w \in \{0,1\}^n$ for each w.

A formal proof that global concavity implies B4 can be found in Ui (2008); the latter refers to our condition as the larger midpoint property. Global concavity imposes nontrivial restrictions on the cross-effects of multivariate functions and, in our model, the support of unobservables and explanatory variables. In particular, if the unobservables and explanatory variables can take values on the entire real line, then the interaction effects must be identically zero for concavity to hold globally. Thus, this requires B1 to be relaxed. For this reason, we cannot recommend basing identification on global concavity explicitly. However, assumption B4 itself can be thought of as a weaker (local) version of discrete concavity that is compatible with our other restrictions.

D.2 *n-Goods bundle model*

Consider an agent who decides whether to buy each of n possible goods. The consumer selects the combination of goods $a = (a_1, a_2, \dots, a_n) \in \{0, 1\}^n$ that maximizes

$$U(a, W, p, \varepsilon, \eta) = \sum_{i \leq n} \big(u_i(W) - p_i + \varepsilon_i\big) \cdot a_i + v\big(S(a), W\big).$$

This model allows the signs of the interaction effects to depend on the identity of all the acquired goods. The model is identical to the previous model if we treat prices as the exclusion restrictions.

Corollary D.1. *If $Z = (-p_i)_{i \leq n}$, identification of the bundles model follows from Theorem* D.1.

D.3 *n-Players potential game*

Consider an extension of the game in Section 2.3 to $\mathcal{N} = \{1, \dots, n\}$ players. Each player $i \in \mathcal{N}$ chooses an action $a_i \in \{0, 1\}$. The payoff of player i from choosing action 1 is

$$U_{1,i}(a_{-i}, W, Z, \varepsilon_i) = u_i(W) + Z_i + \varepsilon_i + v_i(a_{-i}, W), \tag{A.3}$$

while the return from action 0, $U_{0,i}(a_{-i}, W, Z, \varepsilon_i)$, is normalized to 0. In addition, we normalize $v_i(a_{-i}, W)$ to 0 when the actions of all players but i are 0. We denote this game by $\Gamma(W, Z, \varepsilon)$. The definition of a pure strategy Nash equilibrium $a^* = (a_i^*)_{i \leq n}$ naturally extends from the two-player case. The same conditions that facilitate identification of the game guarantee that its equilibrium set (in pure strategies), $\mathcal{D}(w, z, \varepsilon)$, is nonempty.

We want to recover $((u_i, v_i)_{i \leq n}, F_{\varepsilon | W, Z})$ from the available distribution of equilibrium choices $\Pr(a \mid w, z)$. We next show that by restricting attention to the class of potential games and relying on the equilibrium selection rule based on potential maximizers, then identification of the game is mathematically equivalent to identification of the model studied at the beginning of this appendix. The theoretical arguments of Ui (2001) and most of the laboratory evidence cited for why players might coordinate on the potential maximizer in Section 2.3 are not specific to two-player games.

A function $U:\{0,1\}^n \times \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}$ is a *potential function* for $\Gamma(w,z,\varepsilon)$ if, for each $i \le n$ and all $a_{-i} \in \{0,1\}^{n-1}$,

$$U(a_i = 1, a_{-i}, w, z, \varepsilon) - U(a_i = 0, a_{-i}, w, z, \varepsilon) = U_{1,i}(a_{-i}, w, z, \varepsilon_i).$$

The function $\Gamma(w,z,\varepsilon)$ is a *potential game* if it admits a potential function. Let us write

$$\mathcal{S}(a) = \{S \subseteq S(a) \mid |S| \ge 2\} \quad \text{and} \quad \mathcal{S}(a,i) = \{S \subseteq S(a) \mid |S| \ge 2, i \in S\},$$

where $\mathcal{S}(a)$ is the set of players who select action 1 in the action profile a. The notation $\mathcal{S}(a,i)$ returns the empty set when $a_i = 0$. Ui (2000, Theorem 3) shows that $\Gamma(w,z,\varepsilon)$ is a potential game if and only if there exists a function

$$\{\tilde{v}(S,w,z) \mid \tilde{v}(S,w,z) : \mathcal{N} \times \mathbb{R}^k \to \mathbb{R}, |S| \ge 2\}$$

such that, for each $i \in \mathcal{N}$ and all $a \in \{0,1\}^n$,

$$U_{1,i}(a,w,z,\varepsilon_i) = u_i(w) + z_i + \varepsilon_i + \sum_{S \in \mathcal{S}(a,i)} \tilde{v}(S,w,z).$$

A potential function is given by

$$U(a,w,z,\varepsilon) = \sum_{i \in \mathcal{N}} \big(u_i(w) + z_i + \varepsilon_i\big) \cdot a_i + v\big(\mathcal{S}(a), w, z\big)$$
$$\text{with } v\big(\mathcal{S}(a), w, z\big) = \sum_{S \in \mathcal{S}(a)} \tilde{v}(S,w,z).$$

Briefly, an n-player game admits a potential representation if the interaction terms are groupwise symmetric. We next present our main result.

COROLLARY D.2. *If a binary game admits a potential function and players coordinate on the potential maximizer, then identification of the game follows from Theorem D.1.*

By Theorem D.1, $((u_i)_{i \le n}, v, F_{\varepsilon|W,Z})$ is identified in our initial framework and it is readily verified that we can recover $\tilde{v}$ from v by a one-to-one change of notation. The four sufficient conditions provided above for the notion of local concavity in condition B4 have counterparts for the analysis of the game. That is, the analog to B4 holds in the game context if one or more of the following conditions are satisfied: there are only two players; the game is of strategic substitutes; the game is of strategic complements and there is a leading player who fulfills the required conditions; the potential function is discrete globally concave.

D.4 *Proof of Theorem* D.1

We first show that if B1–B4 are satisfied, then $F_{\varepsilon|W}$ is identified. We then show that if B1–B3 are satisfied and $F_{\varepsilon|W}$ is identified, then $((u_i)_{i \le n}, v)$ is also identified.

Identification of $F_{\varepsilon|W}$. By B4, for each $W = w$, there exists a known vector $\widehat{a}^w \in \{0,1\}^n$ such that, for any w, z and ε, $U(a, w, z, \varepsilon)$ is maximized at $\widehat{a}^w$ if $U(\widehat{a}_i^w, \widehat{a}_{-i}^w, w, z, \varepsilon) \geq U(a_i', \widehat{a}_{-i}^w, w, z, \varepsilon)$ for all $i \in \mathcal{N}$ and $a_i' = 1 - \widehat{a}_i^w$. This condition holds if, for all $i \in \mathcal{N}$,

$$\begin{aligned}&\big(1(\widehat{a}_i^w = 1) - 1(\widehat{a}_i^w = 0)\big)\varepsilon_i \\ &\quad\geq v\big(S(a'), w\big) - v\big(S(\widehat{a}^w), w\big) - \big(1(\widehat{a}_i^w = 1) - 1(\widehat{a}_i^w = 0)\big)\big(u_i(w) + z_i\big),\end{aligned} \tag{A.4}$$

where a' is obtained from $\widehat{a}^w$ by changing only $\widehat{a}_i^w$ and $1(\cdot)$ is the usual indicator function.

We next recover $F_{\varepsilon|W}$ from variation in z using $\Pr(\widehat{a}^w \mid w, z)$. From B4 we get that

$$\Pr\big(\widehat{a}^w \mid w, z\big) = \Pr\big(\text{(A.4) holds for all } i \in \mathcal{N} \mid w, z\big).$$

Define the random variable μ_i for each $i \in \mathcal{N}$ to be

$$\begin{aligned}\mu_i = &\big(1(\widehat{a}_i^w = 1) - 1(\widehat{a}_i^w = 0)\big)\varepsilon_i \\ &- \big(v\big(S(a'), w\big) - v\big(S(\widehat{a}^w), w\big)\big) + \big(1(\widehat{a}_i^w = 1) - 1(\widehat{a}_i^w = 0)\big)\big(u_i(w)\big).\end{aligned}$$

Let $\mu \equiv (\mu_1, \ldots, \mu_n)$, which is independent of Z conditional on $W = w$. Therefore,

$$\Pr\big(\widehat{a}^w \mid w, z\big) = \Pr\big(\mu_i \geq -\big(1(\widehat{a}_i^w = 1) - 1(\widehat{a}_i^w = 0)\big)z_i \text{ for all } i \in \mathcal{N} \mid w, z\big).$$

We identify the upper probabilities of the vector μ, conditional on w, at all points

$$\tilde{z} = \big(-\big(1(\widehat{a}_1^w = 1) - 1(\widehat{a}_1^w = 0)\big)z_1, \ldots, -\big(1(\widehat{a}_n^w = 1) - 1(\widehat{a}_n^w = 0)\big)z_n\big).$$

By the large support in B1 and the fact that $\tilde{z}$ is at most a sign change from z, the random vector $\tilde{Z}$, defined in the obvious way, has support on all of $\mathbb{R}^n$. Therefore, we learn the upper tail probabilities of μ conditional on $W = w$ for all points of evaluation μ^*. Upper tail probabilities completely determine a random vector's distribution, so we also identify the lower tail probabilities of μ conditional on w, also known as the joint CDF of μ conditional on w. Note that ε_i is the only random variable in μ_i, conditional on $W = w$. By B2(ii), $E(\varepsilon \mid W) = 0$. Therefore, up to the possible sign change in $(1(\widehat{a}_i^w = 1) - 1(\widehat{a}_i^w = 0))$, the distribution of ε conditional on w is identified from the distribution of $\mu - E(\mu \mid w)$ conditional on w.

Identification of $((u_i)_{i \leq n}, v)$. The remaining argument conditions on $W = w$. Recall the utility function in equation (A.2). The deterministic portion of utility plays a key role in the identification argument. Therefore, let

$$Q(a, w, Z) \equiv \sum_{i \in \mathcal{N}} \big(u_i(w) + Z_i\big) \cdot a_i + v\big(S(a), w\big).$$

Our location normalization is that $Q(a, w, Z) = 0$ for $a = (0, 0, \ldots, 0)$.

For expositional ease, we order the elements of $\{0,1\}^n$ in terms of the lexicographic order so that $a^1 = (0, 0, \ldots, 0)$, $a^2 = (1, 0, \ldots, 0)$, $\ldots$, and $a^{2^n} = (1, 1, \ldots, 1)$. In addition, we define

$$\triangle^j \varepsilon(a') \equiv \sum_{i\in\mathcal{N}} \varepsilon_i 1(a^j_i = 1) - \sum_{i\in\mathcal{N}} \varepsilon_i 1(a'_i = 1) \quad \text{and}$$

$$\triangle^j Q(a', w, Z) \equiv Q(a', w, Z) - Q(a^j, w, Z).$$

We indicate by $\triangle Q(a', w, Z)$ and $\triangle\varepsilon(a')$ the $(2^n - 1)$-dimensional vectors

$$\big(Q(a', w, Z) - Q(a^j, w, Z)\big)_{j\le 2^n, a^j\neq a'} \quad \text{and} \quad \Big(\sum_{i\in\mathcal{N}} \varepsilon_i 1(a^j_i = 1) - \sum_{i\in\mathcal{N}} \varepsilon_i 1(a'_i = 1)\Big)_{j\le 2^n, a^j\neq a'}.$$

Given this notation, for each a',

$$P\big(\triangle Q(a', w, z) \mid w, z; F_{\varepsilon|w,z}\big) \equiv \Pr\big(\triangle\varepsilon(a') \le \triangle Q(a', w, z) \mid w, z\big)$$

is the probability of observing the choice vector a' conditional on $W, Z = w, z$; that is, $\Pr(a' \mid w, z)$. More formally, let $F_{\triangle\varepsilon(a')|w,z}$ be the distribution of $\triangle\varepsilon(a')$. Then

$$\begin{aligned} &P\big(\triangle Q(a', w, z) \mid w, z; F_{\varepsilon|w,z}\big) \\ &\quad\equiv \int\cdots\int 1\big(\triangle^1\varepsilon(a') \le \triangle^1 Q(a', w, z)\big)\cdots 1\big(\triangle^{2^n}\varepsilon(a') \le \triangle^{2^n} Q(a', w, z)\big)\, dF_{\triangle\varepsilon(a')|w,z}. \end{aligned}$$

The researcher can identify $\Pr(a' \mid w, z)$ directly from the data.

Let $\tilde{Q}(a, w, Z) \neq Q(a, w, Z)$. As Z enters both $\tilde{Q}$ and Q in the same way, this means that one or more of $(u_i(w))_{i\in\mathcal{N}}$ and $v(S(a), w)$ differ across $\tilde{Q}$ and Q for $W = w$. We next show that, without loss of generality, $P(\triangle Q(a', w, z) \mid w, z; F_{\varepsilon|w,z}) > P(\triangle\widetilde{Q}(a', w, z) \mid w, z; F_{\varepsilon|w,z})$ for some a' and $Z = z$, so that Q is identified and we can then recover $((u_i)_{i\le n}, v)$. To this end, let

$$C(w, z) \equiv \arg\max_a \big\{\big(Q(a, w, z) - \tilde{Q}(a, w, z)\big) \mid a \in \{0, 1\}^n\big\}.$$

Note that by the formulas for $(Q(a, w, z) - \widetilde{Q}(a, w, z))$, $C(w, z)$ does not vary with z. Also suppose $\max_a(Q(a, w, z) - \widetilde{Q}(a, w, z)) > 0$; the other case follows by a similar argument. Define $D(w, z) \equiv \{a \notin C(w, z) \mid a \in \{0, 1\}^n\}$. We know $C(w, z) \neq \emptyset$. The fact that $D(w, z) \neq \emptyset$ follows as

$$Q\big(a = (0, 0, \dots, 0), w, z\big) = \widetilde{Q}\big(a = (0, 0, \dots, 0), w, z\big) = 0 \tag{A.5}$$

and we supposed that $\max_a(Q(a, w, z) - \widetilde{Q}(a, w, z)) > 0$.

Fix some $a' \in C(w, z)$. We know that $Q(a', w, z) - \widetilde{Q}(a', w, z) = Q(a, w, z) - \widetilde{Q}(a, w, z)$ for all $a \in C(w, z)$, and $Q(a', w, z) - \widetilde{Q}(a', w, z) > Q(a, w, z) - \widetilde{Q}(a, w, z)$ for all $a \in D(w, z)$. Rearranging terms,

$$Q\big(a', w, z\big) - Q(a, w, z) = \widetilde{Q}\big(a', w, z\big) - \widetilde{Q}(a, w, z) \quad \text{for all } a \in C(w, z),$$

$$Q\big(a', w, z\big) - Q(a, w, z) > \widetilde{Q}\big(a', w, z\big) - \widetilde{Q}(a, w, z) \quad \text{for all } a \in D(w, z).$$

By B2 and B3, the argument in the following paragraphs ensures that we can find z such that

$$\int \cdots \int 1\big(\triangle^1 \varepsilon(a') \leq \triangle^1 Q(a', w, z)\big) \cdots 1\big(\triangle^{2^n} \varepsilon(a') \leq \triangle^{2^n} Q(a', w, z)\big)\, dF_{\triangle\varepsilon(a')|w}$$
$$> \int \cdots \int 1\big(\triangle^1 \varepsilon(a') \leq \triangle^1 \widetilde{Q}(a', w, z)\big) \cdots 1\big(\triangle^{2^n} \varepsilon(a') \leq \triangle^{2^n} \widetilde{Q}(a', w, z)\big)\, dF_{\triangle\varepsilon(a')|w},$$

that is, $P(\triangle Q(a', w, z) \mid w, z; F_{\varepsilon|w}) > P(\triangle \widetilde{Q}(a', w, z) \mid w, z; F_{\varepsilon|w})$. Therefore, Q is nonconstructively identified at w, and hence $((u_i(w))_{i\leq n}, v(w))$ is identified as well.

As mentioned previously, we need to find an appropriate value for z. This choice of z involves an additional detail that we address next. The inequalities that allow us to show that $P(\triangle Q(a', w, z) \mid w, z; F_{\varepsilon|w}) > P(\triangle \widetilde{Q}(a', w, z) \mid w, z; F_{\varepsilon|w})$ are the ones that involve $a \in D(w, z)$. Some of the inequalities involving $a \in D(w, z)$ may be implied by other inequalities for any z. To see this, suppose there are two substitute items and let $a' = (0, 0)$, $C(w, z) = \{(0, 0), (1, 0), (0, 1)\}$, and $D(w, z) = \{(1, 1)\}$. Under substitutes, $Q((0, 0), w, z) \geq Q((1, 0), w, z)$ and $Q((0, 0), w, z) \geq Q((0, 1), w, z)$ together imply $Q((0, 0), w, z) > Q((1, 1), w, z)$. In this example, the fact that two inequalities imply a third means that marginal changes in the interaction term, v, will not affect the probability of the outcome $(0, 0)$. Therefore, $a' = (0, 0)$ does not allow us to effectively distinguish Q from $\widetilde{Q}$ as we just claimed. Notice that we assumed $(0, 0) \in C(w, z)$, which is not possible as we are covering the case $\max_a(Q(a, w, z) - \widetilde{Q}(a, w, z)) > 0$ and (A.5) holds. The next argument extends this idea.

We now show that there always exists some $a'' \in D(w, z)$ for which the inequality $Q(a', w, z) \geq Q(a'', w, z)$ is not directly implied from $Q(a', w, z) \geq Q(a, w, z)$ for all $a \neq a', a''$. By contradiction, as we illustrated above, we will show that if this were not true, then $(0, 0, \dots, 0) \in C(w, z)$ which is not possible as $\max_a(Q(a, w, z) - \widetilde{Q}(a, w, z)) > 0$ and (A.5) holds.

For each $a' \in C(w, z)$, there are (at least) n inequalities that are not implied by the others. These inequalities correspond to vectors of actions that differ from a' regarding one single item. To see this, let a'' be equal to a' except for some item i that is chosen at a' but not at a''. Then $Q(a', w, z) \geq Q(a'', w, z)$ if and only if

$$u_i(w) + z_i + v\big(S(a'), w\big) \geq v\big(S(a''), w\big).$$

All other inequalities $Q(a', w, z) \geq Q(a, w, z)$ with $a \neq a', a''$ will involve at least an excluded variable z_j with $j \neq i$. Thus, we can always find a vector z such that $Q(a', w, z) \geq Q(a, w, z)$ with $a \neq a', a''$ and yet $Q(a', w, z) < Q(a'', w, z)$. Thus, assume $a'' \in C(w, z)$. By repeating this process $\|a'\|$ times (the number of items being selected in bundle a'), we need to assume $(0, 0, \dots, 0) \in C(w, z)$. But this is not possible, as we explained before. □

References

Manski, C. (1988), "Identification of binary response models." *Journal of Econometrics*, 83, 729–738. [2]

Topkis, D. (1998), *Supermodularity and Complementarity*. Princeton University Press, Princeton. [1]

Ui, T. (2000), "A Shapley value representation of potential games." *Games Economic Behavior*, 31, 121–135. [7]

Ui, T. (2001), "Robust equilibria of potential games." *Econometrica*, 69, 1373–1380. [6]

Ui, T. (2008), "Discrete concavity for potential games." *International Game Theory Review*, 10, 137–143. [6]

Co-editor Rosa L. Matzkin handled this manuscript.

Manuscript received 11 September, 2014; final version accepted 9 October, 2016; available online 7 March, 2017.