Econometrica Supplementary Material

SUPPLEMENT TO "NONPARAMETRIC ANALYSIS OF RANDOM UTILITY MODELS"

(*Econometrica*, Vol. 86, No. 6, November 2018, 1883–1909)

YUICHI KITAMURA
Department of Economics, Cowles Foundation, Yale University

JÖRG STOYE
Department of Economics, Cornell University, Department of Economics, University of Bonn, and Hausdorff Center for Mathematics

APPENDIX A: PROOFS AND FURTHER DETAILS OF INFERENTIAL PROCEDURES

PROOF OF THEOREM 3.1: THE PROOF USES nonstochastic demand systems, which can be identified with vectors $(d_1, \dots, d_J) \in \mathcal{B}_1 \times \cdots \times \mathcal{B}_J$. Such a system is rationalizable if $d_j \in \arg\max_{y \in \mathcal{B}_j} u(y)$, $j = 1, \dots, J$, for some utility function u.

Rationalizability of nonstochastic demand systems is well understood. In particular, and irrespective of whether we define rationalizability by GARP or SARP, it is decidable from knowing the preferences directly revealed by choices, hence from knowing patches containing $(d_1, \dots, d_J)$. It follows that for all nonstochastic demand systems that select from the same patches, either all or none are rationalizable.

Fix $(P_1, \dots, P_J)$. Let the set $\mathcal{Y}^*$ collect one "representative" element (e.g., the geometric center point) of each patch. Let $(P_1^*, \dots, P_J^*)$ be the unique stochastic demand system concentrated on $\mathcal{Y}^*$ and having the same vector representation as $(P_1, \dots, P_J)$. The previous paragraph established that demand systems can be arbitrarily perturbed within patches, so $(P_1, \dots, P_J)$ is rationalizable iff $(P_1^*, \dots, P_J^*)$ is. It follows that rationalizability of $(P_1, \dots, P_J)$ can be decided from its vector representation π, and that it suffices to analyze stochastic demand systems supported on $\mathcal{Y}^*$.

Now, any stochastic demand system is rationalizable iff it is a mixture of rationalizable nonstochastic systems. Since $\mathcal{Y}^*$ is finite, there are finitely many nonstochastic demand systems supported on it; of these, a subset will be rationalizable. Noting that these demand systems are characterized by binary vector representations corresponding to columns of A, the statement of the theorem is immediate for the restricted class of stochastic demand systems supported on $\mathcal{Y}^*$. *Q.E.D.*

PROOF OF THEOREM 3.2: We begin with some preliminary observations. Throughout this proof, $c(\mathcal{B}_i)$ denotes the object actually chosen from budget $\mathcal{B}_i$.

(i) If there is a choice cycle of any finite length, then there is a cycle of length 2 or 3 (where a cycle of length 2 is a WARP violation). To see this, assume there exists a length N choice cycle $c(\mathcal{B}_i) \succ c(\mathcal{B}_j) \succ c(\mathcal{B}_k) \succ \cdots \succ c(\mathcal{B}_i)$. If $c(\mathcal{B}_k) \succ c(\mathcal{B}_i)$, then a length 3 cycle has been discovered. Else, there exists a length $N-1$ choice cycle $c(\mathcal{B}_i) \succ c(\mathcal{B}_k) \succ \cdots \succ c(\mathcal{B}_i)$. The argument can be iterated until $N = 4$.

(ii) Call a length 3 choice cycle *irreducible* if it does not contain a length 2 cycle. Then a choice pattern is rationalizable iff it contains no length 2 cycles and also no irreducible length 3 cycles. (In particular, one can ignore reducible length 3 cycles.) This follows trivially from (i).

Yuichi Kitamura: yuichi.kitamura@yale.edu
Jörg Stoye: stoye@cornell.edu

 https://doi.org/10.3982/ECTA14478

(iii) Let $J = 3$ and $M = 1$, that is, assume there are three budgets but two of them fail to intersect. Then any length 3 cycle is reducible. To see this, assume w.l.o.g. that $\mathcal{B}_1$ is below $\mathcal{B}_3$; thus $c(\mathcal{B}_3) \succ c(\mathcal{B}_1)$ by monotonicity. If there is a choice cycle, we must have $c(\mathcal{B}_1) \succ c(\mathcal{B}_2) \succ c(\mathcal{B}_3)$. Cycle $c(\mathcal{B}_1) \succ c(\mathcal{B}_2)$ implies that $c(\mathcal{B}_2)$ is below $\mathcal{B}_1$; thus it is below $\mathcal{B}_3$. Cycle $c(\mathcal{B}_2) \succ c(\mathcal{B}_3)$ implies that $c(\mathcal{B}_3)$ is below $\mathcal{B}_2$. Thus, choice from $(\mathcal{B}_2, \mathcal{B}_3)$ violates WARP.

We are now ready to prove the main result. The nontrivial direction is "only if": thus it suffices to show that if choice from $(\mathcal{B}_1, \dots, \mathcal{B}_{J-1})$ is rationalizable but choice from $(\mathcal{B}_1, \dots, \mathcal{B}_J)$ is not, then choice from $(\mathcal{B}_{M+1}, \dots, \mathcal{B}_J)$ cannot be rationalizable. By observation (ii), if $(\mathcal{B}_1, \dots, \mathcal{B}_J)$ is not rationalizable, it contains either a 2-cycle or an irreducible 3-cycle. Because choice from all triplets within $(\mathcal{B}_1, \dots, \mathcal{B}_{J-1})$ is rationalizable by assumption, it is either the case that some $(\mathcal{B}_i, \mathcal{B}_J)$ constitute a 2-cycle or that some triplet $(\mathcal{B}_i, \mathcal{B}_k, \mathcal{B}_J)$, where $i < k$ w.l.o.g., reveals an irreducible choice cycle. In the former case, $\mathcal{B}_i$ must intersect $\mathcal{B}_J$; hence, $i > M$ and, hence, the conclusion. In the latter case, if $k \leq M$, the choice cycle must be a 2-cycle in $(\mathcal{B}_i, \mathcal{B}_k)$, contradicting rationalizability of $(\mathcal{B}_1, \dots, \mathcal{B}_{J-1})$. If $i \leq M$, the choice cycle is reducible by (iii). Thus, $i > M$; hence, the conclusion. *Q.E.D.*

PROOF OF LEMMA 4.1: Letting $\nu_\tau = \nu - (\tau/H)\mathbf{1}_H$ in $\mathcal{C}_\tau = \{A\nu | \nu \geq (\tau/H)\mathbf{1}_H\}$, we have

$$\begin{aligned}
\mathcal{C}_\tau &= \{A[\nu_\tau + (\tau/H)\mathbf{1}_H] | \nu_\tau \geq 0\} \\
&= \mathcal{C} \oplus (\tau/H)A\mathbf{1}_H \\
&= \{t : t - (\tau/H)A\mathbf{1}_H \in \mathcal{C}\},
\end{aligned}$$

where $\oplus$ signifies Minkowski sum. Define

$$\phi = -BA\mathbf{1}_H/H.$$

Using the $\mathcal{H}$ representation of $\mathcal{C}$,

$$\begin{aligned}
\mathcal{C}_\tau &= \{t : B(t - (\tau/H)A\mathbf{1}_H) \leq 0\} \\
&= \{t : Bt \leq -\tau\phi\}.
\end{aligned}$$

Note that the above definition of ϕ implies $\phi \in \mathrm{col}(B)$. Also define

$$\begin{aligned}
\Phi &:= -BA \\
&= -\begin{bmatrix} b_1' \\ \vdots \\ b_m' \end{bmatrix} [a_1, \dots, a_H] \\
&= \{\Phi_{kh}\},
\end{aligned}$$

where $\Phi_{kh} = b_k' a_h$, $1 \leq k \leq m$, $1 \leq h \leq H$, and let e_h be the hth standard unit vector in $\mathbf{R}^H$. Since $e_h \geq 0$, the $\mathcal{V}$ representation of $\mathcal{C}$ implies that $Ae_h \in \mathcal{C}$ and, thus,

$$BAe_h \leq 0$$

by its $\mathcal{H}$ representation. Therefore,

$$\Phi_{kh} = -e_k' BAe_h \geq 0, \quad 1 \leq k \leq m, 1 \leq h \leq H. \tag{S.1}$$

But if $k \le \bar{m}$, it cannot be that

$$a_j \in \{x : b_k' x = 0\} \quad \text{for all } j,$$

whereas

$$b_k' a_h = 0$$

holds for $\bar{m} + 1 \le k \le m$, $1 \le h \le H$. Therefore, if $k \le \bar{m}$, $\Phi_{kh} = b_k' a_h$ is nonzero at least for one h, $1 \le h \le H$, whereas if $k > \bar{m}$, $\Phi_{kh} = 0$ for every h. Since (S.1) implies that all of $\{\Phi_{kh}\}_{h=1}^H$ are nonnegative, we conclude that

$$\phi_k = \frac{1}{H} \sum_{h=1}^{H} \Phi_{kh} > 0$$

for every $k \le \bar{m}$ and $\phi_k = 0$ for every $k > \bar{m}$. We now have

$$\mathcal{C}_\tau = \{t : Bt \le -\tau\phi\},$$

where ϕ satisfies the stated properties (i) and (ii). *Q.E.D.*

Before we present the proof of Theorem 4.2, it is necessary to specify a class of distributions, to which we impose a mild condition that guarantees stable behavior of the statistic $\mathcal{J}_N$. To this end, we further specify the nature of each row of B. Recall that w.l.o.g., the first $\bar{m}$ rows of B correspond to inequality constraints, whereas the rest of the rows represent equalities. Note that the $\bar{m}$ inequalities include nonnegativity constraints $\pi_{i|j} \ge 0$, $1 \le i \le I_j$, $1 \le j \le J$, represented by the row of B consisting of a negative constant for the corresponding element and zeros otherwise. Likewise, the identities that $\sum_{i=1}^{I_j} \pi_{i|j}$ is constant across $1 \le j \le J$ are included in the set of equality constraints.[1] We show in the proof that the presence of these "definitional" equalities/inequalities, which always hold by construction of $\hat{\pi}$, do not affect the asymptotic theory even when they are (close to) binding. Define $\mathcal{K} = \{1, \dots, m\}$, and let $\mathcal{K}^D$ be the set of indices for the rows of B corresponding to the above nonnegativity constraints and the constant-sum constraints. Let $\mathcal{K}^R = \mathcal{K} \setminus \mathcal{K}^D$, so that $b_k' \pi \le 0$ represents an economic restriction if $k \in \mathcal{K}^R$.[2] Recalling that the choice vectors $(d_{j|1}, \dots, d_{j|N_j})$ are independent and identically distributed (IID) within each time period j, $1 \le j \le J$, let d_j denote the choice vector of a consumer facing budget j (therefore w.l.o.g we can let $d_j = d_{j|1}$). Define $d = [d_1', \dots, d_J']'$, a random I vector of binary variables. Note $\mathrm{E}[d] = \pi$. Let

$$\begin{aligned} g &= Bd \\ &= [g_1, \dots, g_m]'. \end{aligned}$$

With these definitions, consider the following requirement.

CONDITION S.1: For each $k \in \mathcal{K}^R$, $\mathrm{var}(g_k) > 0$ and $\mathrm{E}[|g_k/\sqrt{\mathrm{var}(g_k)}|^{2+c_1}] < c_2$ hold, where c_1 and c_2 are positive constants.

[1]If we impose the (redundant) restriction $\mathbf{1}_H' \nu = 1$ in the definition of $\mathcal{C}$, then the corresponding equality restrictions would be $\sum_{i=1}^{I_j} \pi_{i|j} = 1$ for every j.

[2]In (4.6), $\mathcal{K}^R$ contains only the last row of the matrix.

This type of condition is standard in the literature; see, for example, Andrews and Soares (2010).

PROOF OF THEOREM 4.2: By applying the Minkowski–Weyl theorem and Lemma 4.1 to $\mathcal{J}_N$ and $\tilde{J}_N(\tau_N)$, we see that our procedure is equivalent to comparing

$$\mathcal{J}_N = \min_{t \in \mathbf{R}^I : Bt \leq 0} N[\hat{\pi} - t]' \Omega [\hat{\pi} - t]$$

to the $1 - \alpha$ quantile of the distribution of

$$\tilde{\mathcal{J}}_N = \min_{t \in \mathbf{R}^I : Bt \leq -\tau_N \phi} N[\tilde{\eta}_{\tau_N} - t]' \Omega [\tilde{\eta}_{\tau_N} - t]$$

with $\phi = [\bar{\phi}', (0, \dots, 0)']'$, $\bar{\phi} \in \mathbf{R}^{\bar{m}}_{++}$, where

$$\tilde{\eta}_{\tau_N} = \hat{\eta}_{\tau_N} + \frac{1}{\sqrt{N}} N(0, \hat{S}),$$
$$\hat{\eta}_{\tau_N} = \underset{t \in \mathbf{R}^I : Bt \leq -\tau_N \phi}{\operatorname{argmin}} N[\hat{\pi} - t]' \Omega [\hat{\pi} - t].$$

Suppose B has m rows and $\operatorname{rank}(B) = \ell$. Define an $\ell \times m$ matrix K such that KB is a matrix whose rows consist of a basis of the row space $\operatorname{row}(B)$. Also let M be an $(I - \ell) \times I$ matrix whose rows form an orthonormal basis of $\ker B = \ker(KB)$, and define $P = \binom{KB}{M}$. Finally, let $\hat{g} = B\hat{\pi}$ and $\hat{h} = M\hat{\pi}$. Then

$$\begin{aligned}
\mathcal{J}_N &= \min_{Bt \leq 0} N \left[\begin{pmatrix} KB \\ M \end{pmatrix} (\hat{\pi} - t) \right]' P^{-1\prime} \Omega P^{-1} \left[\begin{pmatrix} KB \\ M \end{pmatrix} (\hat{\pi} - t) \right] \\
&= \min_{Bt \leq 0} N \begin{pmatrix} K[\hat{g} - Bt] \\ \hat{h} - Mt \end{pmatrix}' P^{-1\prime} \Omega P^{-1} \begin{pmatrix} K[\hat{g} - Bt] \\ \hat{h} - Mt \end{pmatrix}.
\end{aligned}$$

Let

$$\mathcal{U}_1 = \left\{ \begin{pmatrix} K\gamma \\ h \end{pmatrix} : \gamma = Bt, h = Mt, B^{\leq} t \leq 0, B^{=} t = 0, t \in \mathbf{R}^I \right\}.$$

Then writing $\alpha = KBt$ and $h = Mt$,

$$\mathcal{J}_N = \min_{\binom{\alpha}{h} \in \mathcal{U}_1} N \begin{pmatrix} K\hat{g} - \alpha \\ \hat{h} - h \end{pmatrix}' P^{-1\prime} \Omega P^{-1} \begin{pmatrix} K\hat{g} - \alpha \\ \hat{h} - h \end{pmatrix}.$$

Also define

$$\mathcal{U}_2 = \left\{ \begin{pmatrix} K\gamma \\ h \end{pmatrix} : \gamma = \begin{pmatrix} \gamma^{\leq} \\ \gamma^{=} \end{pmatrix}, \gamma^{\leq} \in \mathbf{R}^{\bar{m}}_{+}, \gamma^{=} = 0, \gamma \in \operatorname{col}(B), h \in \mathbf{R}^{I - \ell} \right\},$$

where $\operatorname{col}(B)$ denotes the column space of B. Obviously $\mathcal{U}_1 \subset \mathcal{U}_2$. Moreover, $\mathcal{U}_2 \subset \mathcal{U}_1$ holds. To see this, let $\binom{K\gamma^*}{h^*}$ be an arbitrary element of $\mathcal{U}_2$. We can always find $t^* \in \mathbf{R}^I$ such that $\gamma^* = Bt^*$. Define

$$t^{**} := t^* + M'h^* - M'Mt^*.$$

Then $Bt^{**} = Bt^* = \gamma^*$; therefore, $B^{\leq}t^{**} \leq 0$ and $B^{=}t^{**} = 0$. Also, $Mt^{**} = Mt^* + MM'h^* - MM'Mt^* = h^*$: therefore, $\binom{K\gamma^*}{h^*}$ is an element of $\mathcal{U}_1$ as well. Consequently,

$$\mathcal{U}_1 = \mathcal{U}_2.$$

We now have

$$\begin{aligned} \mathcal{J}_N &= \min_{\binom{\alpha}{h} \in \mathcal{U}_2} N \begin{pmatrix} K\hat{g} - \alpha \\ \hat{h} - h \end{pmatrix}' P^{-1\prime} \Omega P^{-1} \begin{pmatrix} K\hat{g} - \alpha \\ \hat{h} - h \end{pmatrix} \\ &= N \min_{\binom{\alpha}{y} \in \mathcal{U}_2} \begin{pmatrix} K\hat{g} - \alpha \\ y \end{pmatrix}' P^{-1\prime} \Omega P^{-1} \begin{pmatrix} K\hat{g} - \alpha \\ y \end{pmatrix}. \end{aligned}$$

Define

$$T(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}' P^{-1\prime} \Omega P^{-1} \begin{pmatrix} x \\ y \end{pmatrix}, \quad x \in \mathbf{R}^{\ell}, y \in \mathbf{R}^{I-\ell},$$

and

$$t(x) := \min_{y \in \mathbf{R}^{I-\ell}} T(x, y), \qquad s(g) := \min_{\gamma = [\gamma^{\leq\prime}, \gamma^{=\prime}]', \gamma^{\leq} \leq 0, \gamma^{=} = 0, \gamma \in \mathrm{col}(B)} t\big(K[g - \gamma]\big).$$

It is easy to see that $t : \mathbf{R}^{\ell} \to \mathbf{R}_+$ is a positive definite quadratic form. We can write

$$\begin{aligned} \mathcal{J}_N &= N \min_{\gamma = [\gamma^{\leq\prime}, \gamma^{=\prime}]', \gamma^{\leq} \leq 0, \gamma^{=} = 0, \gamma \in \mathrm{col}(B)} t\big(K[\hat{g} - \gamma]\big) \\ &= N s(\hat{g}) \\ &= s(\sqrt{N}\hat{g}). \end{aligned}$$

We now show that tightening can turn nonbinding inequality constraints into binding ones, but not vice versa. Note that, as will be seen below, this observation uses diagonality of Ω and the specific geometry of the cone $\mathcal{C}$. Let $\hat{\gamma}^k_{\tau_N}$, $\hat{g}^k$, and ϕ^k denote the kth elements of $\hat{\gamma}_{\tau_N} = B\hat{\eta}_{\tau_N}$, $\hat{g}$, and ϕ. Moreover, define $\gamma_\tau(g) = [\gamma^1(g), \dots, \gamma^m(g)]' = \mathrm{argmin}_{\gamma = [\gamma^{\leq\prime}, \gamma^{=\prime}]', \gamma^{\leq} \leq -\tau\bar{\phi}, \gamma^{=} = 0, \gamma \in \mathrm{col}(B)} t(K[g - \gamma])$ for $g \in \mathrm{col}(B)$, and let $\gamma^k_\tau(g)$ be its kth element. Then $\hat{\gamma}_{\tau_N} = \gamma_{\tau_N}(\hat{g})$. Finally, define $\beta_\tau(g) = \gamma_\tau(g) + \tau\phi$ for $\tau > 0$ and let $\beta^k_\tau(g)$ denote its kth element. Note $\gamma^k_\tau(g) = \phi^k = \beta^k_\tau(g) = 0$ for every $k > \bar{m}$ and g. Now we show that for each $k \leq \bar{m}$ and for some $\delta > 0$,

$$\beta^k_\tau(g) = 0$$

if $|g^k| \leq \tau\delta$ and $g^j \leq \tau\delta$, $1 \leq j \leq \bar{m}$. In what follows we first show this for the case with $\Omega = \mathbf{I}_I$, where $\mathbf{I}_I$ denotes the I-dimensional identity matrix; then we generalize the result to the case where Ω can have arbitrary positive diagonal elements.

For $\tau > 0$ and $\delta > 0$, define hyperplanes

$$\begin{aligned} H^\tau_k &= \big\{x : b'_k x = -\tau\phi^k\big\}, \\ H_k &= \big\{x : b'_k x = 0\big\}, \end{aligned}$$

half-spaces

$$H^\tau_{\angle k}(\delta) = \big\{x : b'_k x \leq \tau\delta\big\},$$

and also

$$S_k(\delta) = \{x \in \mathcal{C} : |b_k' x| \leq \tau\delta\}$$

for $1 \leq k \leq m$. Define

$$L = \bigcap_{k=\bar{m}+1}^{m} H_k,$$

a linear subspace of $\mathbf{R}^I$. In what follows we show that for small enough $\delta > 0$, every element $x^* \in \mathbf{R}^I$ such that

$$x^* \in S_1(\delta) \cap \cdots \cap S_q(\delta) \cap H^\tau_{\angle q+1}(\delta) \cap \cdots \cap H^\tau_{\angle m}(\delta) \quad \text{for some } q \in \{1, \ldots, \bar{m}\} \tag{S.2}$$

satisfies

$$x^*|\mathcal{C}_\tau \in H_1^\tau \cap \cdots \cap H_q^\tau \cap L, \tag{S.3}$$

where $x^*|\mathcal{C}_\tau$ denotes the orthogonal projection of x^* on $\mathcal{C}_\tau$. Let $g^{*k} = b_k' x^*$, $k = 1, \ldots, m$. Note that an element x^* fulfills (S.2) iff $|g^{*k}| \leq \tau\delta$, $1 \leq k \leq q$, and $g^{*j} \leq \tau\delta$, $q+1 \leq j \leq \bar{m}$. Likewise, (S.3) holds iff $\beta_k^\tau(g^*) = 0$, $1 \leq k \leq q$ (recall that $\beta_k^\tau(g^*) = 0$ always holds for $k > \bar{m}$). Thus, so as to establish the desired property of the function $\beta_\tau(\cdot)$, we show that (S.2) implies (S.3). Suppose it does not hold; then without loss of generality, for an element x^* that satisfies (S.2) for an arbitrary small $\delta > 0$, we have

$$x^*|\mathcal{C}_\tau \in H_1^\tau \cap \cdots \cap H_r^\tau \cap L \quad \text{and} \quad x^*|\mathcal{C}_\tau \notin H_j^\tau, \quad r+1 \leq j \leq q \tag{S.4}$$

for some $1 \leq r \leq q-1$. Define half-spaces

$$H^\tau_{\angle k} = \{x : b_k' x \leq -\tau\phi^k\},$$
$$H_{\angle k} = \{x : b_k' x \leq 0\}$$

for $1 \leq k \leq m$, $\tau > 0$, and also let

$$F = H_1 \cap \cdots \cap H_r \cap \mathcal{C}.$$

Then for (S.4) to hold for some $x^* \in \mathbf{R}^I$ satisfying (S.2) for an arbitrary small $\delta > 0$, we must have

$$F|(H_1^\tau \cap \cdots \cap H_r^\tau \cap L) \subset \operatorname{int}(H^\tau_{\angle r+1} \cap \cdots \cap H^\tau_{\angle q}).$$

(Recall that the notation | signifies orthogonal projection. Also note that if $\dim(F) = 1$, then (S.4) does not occur under (S.2).) Therefore, if we let

$$\Delta(J) = \{x \in \mathbf{R}^I : \mathbf{1}_I' x = J, x \geq 0\},$$

that is, the simplex with vertices $(J, 0, \ldots, 0), \ldots, (0, \ldots, 0, J)$, we have

$$(F \cap \Delta(J))|(H_1^\tau \cap \cdots \cap H_r^\tau \cap L) \subset \operatorname{int}(H^\tau_{\angle r+1} \cap \cdots \cap H^\tau_{\angle q}). \tag{S.5}$$

Let $\{a_1, \ldots, a_H\} = \mathcal{A}$ denote the collection of the column vectors of A. Then {the vertices of $F \cap \Delta(J)$} $\in \mathcal{A}$. Let $\bar{a}, \bar{\bar{a}} \in F \cap \Delta(J)$. Let $B(\varepsilon, x)$ denote the ε (-open) ball with center

$x \in \mathbf{R}^I$. By (S.5),

$$B\left(\varepsilon, \left(\bar{a}\middle|\bigcap_{j=1}^{r} H_j^\tau \cap L\right)\right) \subset \operatorname{int}\left(H_{\angle r+1}^\tau \cap \cdots \cap H_{\angle q}^\tau\right) \cap H_{\angle 1} \cap \cdots \cap H_{\angle r}$$

holds for small enough $\varepsilon > 0$. Let $\bar{a}^\tau := \bar{a} + \frac{\tau}{H} A \mathbf{1}_H$ and $\bar{\bar{a}}^\tau := \bar{\bar{a}} + \frac{\tau}{H} A \mathbf{1}_H$. Then

$$\begin{aligned}\left(\left(\bar{a}\middle|\left(\bigcap_{j=1}^{r} H_j^\tau\right) \cap L\right) - \bar{a}\right)'(\bar{\bar{a}} - \bar{a}) &= \left(\left(\bar{a}|\left(\bigcap_{j=1}^{r} H_j^\tau\right) \cap L\right) - \bar{a}\right)'\left(\bar{\bar{a}}^\tau - \bar{a}^\tau\right) \\ &= 0\end{aligned}$$

since $\bar{a}^\tau, \bar{\bar{a}}^\tau \in (\bigcap_{j=1}^r H_j^\tau) \cap L$. We can then take $z \in B(\varepsilon, (\bar{a}|(\bigcap_{j=1}^r H_j^\tau) \cap L))$ such that $(z - \bar{a})'(\bar{\bar{a}} - \bar{a}) < 0$. By construction $z \in \mathcal{C}$, which implies the existence of a triplet $(a, \bar{a}, \bar{\bar{a}})$ of distinct elements in $\mathcal{A}$ such that $(a - \bar{a})'(\bar{\bar{a}} - \bar{a}) < 0$. In what follows we show that this cannot happen; then the desired property of β_τ is established.

So let us now show that

$$(a_1 - a_0)'(a_2 - a_0) \geq 0 \quad \text{for every triplet } (a_0, a_1, a_2) \text{ of distinct elements in } \mathcal{A}. \tag{S.6}$$

Noting that $a_i'a_j$ just counts the number of budgets on which i and j agree, define

$$\phi(a_i, a_j) = J - a_i'a_j,$$

the number of disagreements. Importantly, note that $\phi(a_i, a_j) = \phi(a_j, a_i)$ and that ϕ is a distance (it is the taxicab distance between elements in $\mathcal{A}$, which are all 0-1 vectors). Now

$$\begin{aligned}&(a_1 - a_0)'(a_2 - a_0) \\ &\quad = a_1'a_2 - a_0'a_2 - a_1'a_0 + a_0'a_0 \\ &\quad = J - \phi(a_1, a_2) - \left(J - \phi(a_0, a_2)\right) - \left(J - \phi(a_0, a_1)\right) + J \\ &\quad = \phi(a_0, a_2) + \phi(a_0, a_1) - \phi(a_1, a_2) \geq 0\end{aligned}$$

by the triangle inequality.

Next we treat the case where Ω is not necessarily $\mathbf{I}_I$. Write

$$\Omega = \begin{bmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \omega_2^2 & \dots & 0 \\ & & \ddots & \\ 0 & \dots & 0 & \omega_I^2 \end{bmatrix}.$$

The statistic $\mathcal{J}_N$ in (4.1) can be rewritten, using the square-root matrix $\Omega^{1/2}$,

$$\mathcal{J}_N = \min_{\eta^* = \Omega^{1/2}\eta : \eta \in C} \left[\hat{\pi}^* - \eta^*\right]'\left[\hat{\pi}^* - \eta^*\right]$$

or

$$\mathcal{J}_N = \min_{\eta^* \in C^*} \left[\hat{\pi}^* - \eta^*\right]'\left[\hat{\pi}^* - \eta^*\right],$$

where

$$\mathcal{C}^* = \{\Omega^{1/2}A\nu | \nu \geq 0\}$$
$$= \{A^*\nu | \nu \geq 0\}$$

with

$$A^* = [a_1^*, \dots, a_H^*], \qquad a_h^* = \Omega^{1/2}a_h, \quad 1 \leq h \leq H.$$

Then we can follow our previous argument, replacing a's with a^*'s, and using

$$\Delta^*(J) = \mathrm{conv}([0, \dots, \omega_i, \dots, 0]' \in \mathbf{R}^I, i = 1, \dots, I)$$

instead of the simplex $\Delta(J)$. Finally, we need to verify that the acuteness condition (S.6) holds for $\mathcal{A}^* = \{a_1^*, \dots, a_H^*\}$.

For two I vectors a and b, define a weighted taxicab metric

$$\phi_\Omega(a, b) := \sum_{i=1}^I \omega_i |a_i - b_i|.$$

Then the standard taxicab metric ϕ used above is ϕ_Ω with $\Omega = \mathbf{I}_I$. Moreover, letting $a^* = \Omega^{1/2}a$ and $b^* = \Omega^{1/2}b$, where each of a and b is an I-dimensional 0-1 vector, we have

$$a^{*\prime}b^* = \sum_{i=1}^I \omega_i[1 - |a_i - b_i|] = \bar{\omega} - \phi_\Omega(a, b)$$

with $\bar{\omega} = \sum_{i=1}^I \omega_i$. Then for every triplet (a_0^*, a_1^*, a_2^*) of distinct elements in $\mathcal{A}^*$,

$$\begin{aligned}(a_1^* - a_0^*)'(a_2^* - a_0^*) &= \bar{\omega} - \phi_\Omega(a_1, a_2) - \bar{\omega} + \phi_\Omega(a_0, a_2) \\ &\quad - \bar{\omega} + \phi_\Omega(a_0, a_1) + \bar{\omega} - \phi_\Omega(a_0, a_0) \\ &= \phi_\Omega(a_1, a_2) - \phi_\Omega(a_0, a_2) - \phi_\Omega(a_0, a_1) \\ &\geq 0,\end{aligned}$$

which is the desired acuteness condition. Since $\mathcal{J}_N$ can be written as the minimum of the quadratic form with identity-matrix weighting subject to the cone generated by a^*s, all the previous arguments developed for the case with $\Omega = \mathbf{I}_I$ remain valid.

Defining $\xi \sim \mathrm{N}(0, \hat{S})$ and $\zeta = B\xi$,

$$\begin{aligned}\tilde{\mathcal{J}}_N &\sim \min_{Bt \leq -\tau_N \phi} N\left[\begin{pmatrix} KB \\ M \end{pmatrix}(\hat{\eta}_{\tau_N} + N^{-1/2}\xi - t)\right]' P^{-1\prime}\Omega P^{-1}\left[\begin{pmatrix} KB \\ M \end{pmatrix}(\hat{\eta}_{\tau_N} + N^{-1/2}\xi - t)\right] \\ &= N \min_{\gamma = [\gamma^{\leq\prime}, \gamma^{=\prime}]', \gamma^{\leq} \leq -\tau_N \bar{\phi}, \gamma^{=} = 0, \gamma \in \mathrm{col}(B)} t(K[\hat{\gamma}_{\tau_N} + N^{-1/2}\zeta - \gamma])\end{aligned}$$

conditional on data $\{\{d_{i|j,n}\}_{i=1}^{I_j}\}_{n=1}^{N_j}$, $j = 1, 2, \dots, J$. Moreover, defining $\gamma^\tau = \gamma + \tau_N \phi$ in the above, and using the definitions of $\beta_\tau(\cdot)$ and $s(\cdot)$,

$$\begin{aligned}\tilde{\mathcal{J}}_N &\sim N \min_{\gamma^\tau = [\gamma^{\tau\leq\prime}, \gamma^{\tau=\prime}]', \gamma^{\tau\leq} \leq 0, \gamma^{\tau=} = 0, \gamma^\tau \in \mathrm{col}(B)} t(K[\hat{\gamma}_{\tau_N} + \tau_N \phi + N^{-1/2}\zeta - \gamma^\tau]) \\ &= N \min_{\gamma^\tau = [\gamma^{\tau\leq\prime}, \gamma^{\tau=\prime}]', \gamma^{\tau\leq} \leq 0, \gamma^{\tau=} = 0, \gamma^\tau \in \mathrm{col}(B)} t(K[\gamma_{\tau_N}(\hat{g}) + \tau_N \phi + N^{-1/2}\zeta - \gamma^\tau])\end{aligned}$$

$$
\begin{aligned}
&= N \min_{\gamma^\tau = [\gamma^{\tau \le \prime}, \gamma^{\tau = \prime}]', \gamma^{\tau \le} \le 0, \gamma^{\tau =} = 0, \gamma^\tau \in \mathrm{col}(B)} t\big(K\big[\beta_{\tau_N}(\hat{g}) + N^{-1/2}\zeta - \gamma^\tau\big]\big) \\
&= s\big(N^{1/2}\beta_{\tau_N}(\hat{g}) + \zeta\big).
\end{aligned}
$$

Let $\varphi_N(\xi) := N^{1/2}\beta_{\tau_N}(\tau_N \xi)$ for $\xi = (\xi_1, \dots, \xi_m)' \in \mathrm{col}(B)$. Then from the property of β_τ shown above, its kth element φ_N^k for $k \le \bar{m}$ satisfies

$$
\varphi_N^k(\xi) = 0
$$

if $|\xi^k| \le \delta$ and $\xi^j \le \delta$, $1 \le j \le m$ for large enough N. Note that $\varphi_N^k(\xi) = N^{1/2}\beta_N^k(\tau_N \xi) = 0$ for $k > \bar{m}$. Define $\hat{\xi} := \hat{g}/\tau_N$. Using the definition of φ_N, we write

$$
\tilde{\mathcal{J}}_N \sim s\big(\varphi_N(\hat{\xi}) + \zeta\big). \tag{S.7}
$$

Now we invoke Theorem 1 of Andrews and Soares (2010; AS henceforth). As noted before, the function t is a positive definite quadratic form on $\mathbf{R}^\ell$ and so is its restriction on $\mathrm{col}(B)$. Then their Assumptions 1–3 hold for the function s defined above if signs are adjusted appropriately, as our formulae deal with negativity constraints, whereas AS work with positivity constraints. (Note that Assumption 1(b) does not apply here since we use a fixed weighting matrix.) The function φ_N in (S.7) satisfies the properties of φ that AS used in their proof of Theorem 1. AS impose a set of restrictions on the parameter space (see their Equation (2.2) on page 124). Their condition (2.2)(vii) is a Lyapounov condition for a triangular array CLT. Following AS, consider a sequence of distributions $\pi_N = [\pi_{1N}', \dots, \pi_{JN}']'$, $N = 1, 2, \dots$, in $\mathcal{P} \cap \mathcal{C}$ such that (a) $\sqrt{N}B\pi_N \to h$ for a nonpositive h as $N \to \infty$ and (b) $\mathrm{Cov}_{\pi_N}(\sqrt{N}B\hat{\pi}) \to \Sigma$ as $N \to \infty$, where Σ is positive semidefinite. The Lyapounov condition holds for $b_k'\hat{\pi}$ under π_N for $k \in \mathcal{K}^R$, as Condition S.1 is imposed for $\pi_N \in \mathcal{P}$. We do not impose Condition S.1 for $k \in \mathcal{K}^D$. Note, however, that (i) the equality $b_k'\hat{\pi} \le 0$ holds by construction for every $k \in \mathcal{K}^D$ and, therefore, its behavior does not affect $\mathcal{J}_N$; (ii) if $\mathrm{var}_{\pi_N}(g_k)$ converges to zero for some $k \in \mathcal{K}^D$, then $\sqrt{N}b_k'[\tilde{\eta}_{\tau_N} - \hat{\eta}_{\tau_N}] = o_p(1)$ and, therefore, its contribution to $\tilde{\mathcal{J}}_N$ is asymptotically negligible in the size calculation. The other conditions in AS, namely (2.2)(i)–(vi), hold trivially. Finally, Assumptions GMS2 and GMS4 of AS are concerned with their thresholding parameter κ_N for the kth moment inequality, and by letting $\kappa_N = N^{1/2}\tau_N\phi_k$; the former holds by the condition $\sqrt{N}\tau_N \uparrow \infty$ and the latter holds by $\tau_N \downarrow 0$. Therefore, we conclude

$$
\liminf_{N\to\infty} \inf_{\pi \in \mathcal{P}\cap\mathcal{C}} \Pr\{\mathcal{J}_N \le \hat{c}_{1-\alpha}\} = 1 - \alpha. \qquad \textit{Q.E.D.}
$$

Further Details of the Procedure in Section 5

The setting in this section is as follows. Let $\tilde{p}_j \in \mathbf{R}_{++}^K$ denote the unnormalized price vector, fixed for each period j. Let $(S, \mathcal{S}, P)$ denote the underlying probability space. Since we have repeated cross sections over J periods, write $P = \bigotimes_{j=1}^J P^{(j)}$, a J-fold product measure.

We first develop a smoothing procedure based on a series estimator (see, for example, Newey (1997)) for π to deal with a situation where total expenditure W is continuously distributed, yet exogenous. We need some notation and definitions to formally state the asymptotic theory behind our procedure with smoothing. With exogeneity, we have

$$
\begin{aligned}
p_{i|j}(\underline{w}_j) &= \Pr\big\{D(\tilde{p}_j/w_{n(j)}, u) \in x_{i|j} | w_{n(j)} = \underline{w}_j\big\} \\
&= \Pr\big\{D(\tilde{p}_j/\underline{w}_j, u) \in x_{i|j}, u \sim P_u\big\},
\end{aligned}
$$

where the second equality follows from the exogeneity assumption. Then $\pi_{i|j} = p_{i|j}(\underline{w}_j)$, which is the estimand in what follows. Define $q^K(w) = (q_{1K}(w), \dots, q_{KK}(w))'$, where $q_{jK}(w)$, $j = 1, \dots, K$, are basis functions (e.g., power series or splines) of w. Instead of sample frequency estimators, for each j, $1 \leq j \leq J$, we use

$$\hat{\pi}_{i|j} = q^{K(j)}(\underline{w}_j)'\widehat{Q}^{-}(j) \sum_{n(j)=1}^{N_j} q^{K(j)}(w_{n(j)})d_{i|j,n(j)}/N_j,$$

$$\widehat{Q}(j) = \sum_{n(j)=1}^{N_j} q^{K(j)}(w_{n(j)})q^{K(j)}(w_{n(j)})'/N_j,$$

$$\hat{\pi}_j = (\hat{\pi}_{1|j}, \dots, \hat{\pi}_{I_j|j})',$$

$$\hat{\pi} = \left(\hat{\pi}_1', \dots, \hat{\pi}_J'\right)'$$

to estimate $\pi_{i|j}$, where A^- denotes a symmetric generalized inverse of A and $K(j)$ is the number of basis functions applied to budget $\mathcal{B}_j$. The estimators $\hat{\pi}_{i|j}$ may not take their values in [0, 1]. This does not seem to cause a problem asymptotically, though as in Imbens and Newey (2009), we may (and do, in the application) instead use

$$\hat{\pi}_{i|j} = G\left(q^{K(j)}(\underline{w}_j)'\widehat{Q}^{-}(j) \sum_{n(j)=1}^{N_j} q^{K(j)}(w_{n(j)})d_{i|j,n(j)}/N_j\right),$$

where G denotes the cumulative distribution function (CDF) of Unif(0, 1). Then an appropriate choice of τ_N is $\tau_N = \sqrt{\frac{\log \underline{n}}{\underline{n}}}$ with

$$\underline{n} = \min_j N_j I_j / \operatorname{trace}\left(v_N^{(j)}\right),$$

where $v_N^{(j)}$ is defined below. Strictly speaking, asymptotics with nonparametric smoothing involve bias, and the bootstrap does not solve the problem. A standard procedure is to claim that one used undersmoothing and can hence ignore the bias, and we follow this convention. The bootstrapped test statistic $\tilde{J}_N$ is obtained by applying the same replacements to the formula (4.5), although generating $\tilde{\eta}_{\tau_N}$ requires a slight modification. Let $\hat{\eta}_{\tau_N}(j)$ be the jth block of the vector $\hat{\eta}_{\tau_N}$, and ket $\hat{v}_N^{(j)}$ satisfy $\hat{v}_N^{(j)} {v_N^{(j)}}^{-1} \to_p \mathbf{I}_{I_j}$, where

$$v_N^{(j)} = \left[\mathbf{I}_{I_j} \otimes q^{K(j)}(\underline{w}_j)' Q_N(j)^{-1}\right] \Lambda_N^{(j)} \left[\mathbf{I}_{I_j} \otimes Q_N^{-1}(j) q^{K(j)}(\underline{w}_j)\right]$$

with $Q_N(j) := \mathrm{E}[q^{K(j)}(w_{n(j)})q^{K(j)}(w_{n(j)})']$, $\Lambda_N^{(j)} := \mathrm{E}[\Sigma^{(j)}(w_{n(j)}) \otimes q^{K(j)}(w_{n(j)})q^{K(j)}(w_{n(j)})']$, and $\Sigma^{(j)}(w) := \operatorname{Cov}[d_{j,n(j)} | w_{n(j)} = w]$. Note that $\Sigma^{(j)}(w) = \operatorname{diag}(p^{(j)}(w)) - p^{(j)}(w)p^{(j)}(w)'$, where $p^{(j)}(w) = [p_{1|j}(w), \dots, p_{I_j|j}(w)]'$. For example, one may use

$$\hat{v}_N^{(j)} = \left[\mathbf{I}_{I_j} \otimes q^{K(j)}(\underline{w}_j)' \hat{Q}^{-}(j)\right] \widehat{\Lambda}(j) \left[\mathbf{I}_{I_j} \otimes \hat{Q}^{-}(j) q^{K(j)}(\underline{w}_j)\right]$$

with $\widehat{\Lambda}(j) = \frac{1}{N_j}\sum_{n(j)=1}^{N_j}[\widehat{\Sigma}^{(j)}(w_{n(j)}) \otimes q^{K(j)}(w_{n(j)})q^{K(j)}(w_{n(j)})']$, $\widehat{\Sigma}^{(j)}(w) = \operatorname{diag}(\widehat{p}^{(j)}(w)) - \widehat{p}^{(j)}(w)\widehat{p}^{(j)}(w)'$, $\widehat{p}^{(j)}(w) = [\widehat{p}_{1|j}(w), \dots, \widehat{p}_{I_j|j}(w)]'$, and $\widehat{p}_{i|j}(w) = q^{K(j)}(w)'\widehat{Q}^{-}(j) \times$

$\sum_{n(j)=1}^{N_j} q^{K(j)}(w_{n(j)})d_{i|j,n(j)}/N_j$. We use $\tilde{\eta}_{\tau_N} = (\tilde{\eta}_{\tau_N}(1)', \ldots, \tilde{\eta}_{\tau_N}(J)')'$ for the smoothed version of $\tilde{J}_N$, where $\tilde{\eta}_{\tau_N}(j) := \hat{\eta}_{\tau_N}(j) + \frac{1}{\sqrt{N_j}} N(0, \hat{v}_N^{(j)})$, $j = 1, \ldots, J$.

Noting that $\{\{d_{i|j,n(j)}\}_{i=1}^{I_j}, w_{n(j)}\}_{n(j)=1}^{N_j}$ are IID distributed within each time period j, $1 \leq j \leq J$, let (d_j, w_j) denote the choice–log-expenditure pair of a consumer facing budget j. Let $d = [d_1', \ldots, d_J']'$ and $\mathbf{w} = [w_1, \ldots, w_J]'$, and define $g = Bd = [g_1, \ldots, g_m]'$ as before. Let $\mathcal{W}_j$ denote the support of $w_{n(j)}$. For a symmetric matrix A, $\lambda_{\min}$ signifies its smallest eigenvalue.

CONDITION S.2: There exist positive constants c_1, c_2, δ, and $\zeta(K)$, $K \in \mathbf{N}$, such that the following statements hold:

(i) We have $\pi \in \mathcal{C}$.

(ii) For each $k \in \mathcal{K}^R$, $\operatorname{var}(g_k|\mathbf{w} = (\underline{w_1}, \ldots, \underline{w_J})') \geq s^2(F_1, \ldots, F_J)$ and $\mathrm{E}[(g_k/s(F_1, \ldots, F_J))^4|\mathbf{w} = (\underline{w_1}, \ldots, \underline{w_J})'] < c_1$ hold for every $(\underline{w_1}, \ldots, \underline{w_J}) \in \mathcal{W}_1 \times \cdots \mathcal{W}_J$.

(iii) The inequality $\sup_{w \in \mathcal{W}_j} |p_{i|j}(w) - q^K(w)'\beta_K^{(j)}| \leq c_1 K^{-\delta}$ holds with some K vector $\beta_K^{(j)}$ for every $K \in \mathbf{N}$, $1 \leq i \leq I_j$, $1 \leq j \leq J$.

(iv) Letting $\widetilde{q}^K := C_{K,j} q^K$, $\lambda_{\min}\mathrm{E}[\widetilde{q}^K(w_{n(j)})\widetilde{q}^K(w_{n(j)})'] \geq c_2$ holds for every K and j, where $C_{K,j}$, $K \in \mathbf{N}$, $1 \leq j \leq J$, are constant nonsingular matrices.

(v) For every $K \in \mathbf{N}$, $\max_j \sup_{w \in \mathcal{W}_j} \|\widetilde{q}^K(w)\| \leq c_2 \zeta(K)$.

Condition S.2(ii) is a version of Condition S.1 that accommodates the conditioning by w and series estimation. Conditions S.2(iii)–(v) are standard regularity conditions commonly used in the series regression literature: (iii) imposes a uniform approximation error bound, (iv) avoids singular design (note the existence of the matrices $C_{K,j}$ suffices), and (v) controls the lengths of the series terms used.

The next condition imposes restrictions on tuning parameters.

CONDITION S.3: The terms τ_N and $K(j)$, $j = 1, \ldots, J$, satisfy $\sqrt{N_j} K^{-\delta}(j) \downarrow 0$, $\zeta(K(j))^2 K(j)/N_j \downarrow 0$, $j = 1, \ldots, J$, $\tau_N \downarrow 0$, and $\sqrt{\underline{n}}\tau_N \uparrow \infty$.

PROOF OF THEOREM 5.1: We begin by introducing some notation.

NOTATION: Let $b_{k,i}$, $k = 1, \ldots, m$, $i = 1, \ldots, I$, denote the (k, i) element of B. Then define

$$b_k(j) = [b_{k,N_1+\cdots N_{j-1}+1}, b_{k,N_1+\cdots N_{j-1}+2}, \ldots, b_{k,N_1+\cdots N_j}]'$$

for $1 \leq j \leq J$ and $1 \leq k \leq m$. Let $B^{(j)} := [b_1(j), \ldots, b_m(j)]' \in \mathbf{R}^{m \times I_j}$. For $F \in \mathcal{F}$ and $1 \leq j \leq J$, define

$$p_F^{(j)}(w) := E_F[d_{j,n(j)}|w_{n(j)} = w], \qquad \pi_F^{(j)} = p_F^{(j)}(\underline{w}_j), \qquad \pi_F = \left[\pi_F^{(1)'}, \ldots, \pi_F^{(J)'}\right]',$$

and

$$\Sigma_F^{(j)}(w) := \mathrm{Cov}_F[d_{j,n(j)}|w_{n(j)} = w].$$

Note that $\Sigma_F^{(j)}(w) = \operatorname{diag}(p_F^{(j)}(w)) - p_F^{(j)}(w)p_F^{(j)}(w)'$.

The proof mimics the proof of Theorem 4.2, except for the treatment of $\hat{\pi}$. Instead of the sequence π_N, $N = 1, 2, \ldots$, in $\mathcal{P} \cap \mathcal{C}$, consider a sequence of distributions $F_N =$

$[F_{1N},\dots,F_{JN}]$, $N=1,2,\dots$, in $\mathcal{F}$ such that $\sqrt{N_j/K(j)}B^{(j)}\pi^{(j)}_{F_N}\to h_j$, $h_j\le 0$, $1\le j\le J$ as $N\to\infty$. Define $Q^{(j)}_{F_N}=\mathrm{E}_{F_N}[q^{K(j)}(w_{n(j)})q^{K(j)}(w_{n(j)})']$ and $\Xi^{(j)}_{F_N}=\mathrm{E}_{F_N}[B^{(j)}\Sigma^{(j)}_{F_N}(w_{n(j)})B^{(j)\prime}\otimes q^{K(j)}(w_{n(j)})q^{K(j)}(w_{n(j)})']$, and let

$$V^{(j)}_{F_N}:=\big[\mathbf{I}_m\otimes q^{K(j)}(\underline{w}_j)'{Q^{(j)}_{F_N}}^{-1}\big]\Xi^{(j)}_{F_N}\big[\mathbf{I}_m\otimes {Q^{(j)}_{F_N}}^{-1}q^{K(j)}(\underline{w}_j)\big]$$

and

$$V_{F_N}:=\sum_{j=1}^{J}V^{(j)}_{F_N}.$$

Then by adapting the proof of Theorem 2 in Newey (1997) to the triangle array for the repeated cross section setting, we obtain

$$\sqrt{N}{V_{F_N}}^{-\frac{1}{2}}B[\hat{\pi}-\pi_{F_N}]\overset{F_N}{\rightsquigarrow}N(0,\mathbf{I}_m).$$

The rest is the same as the proof of Theorem 4.2. *Q.E.D.*

Next we turn to the definition of our endogeneity-corrected estimator $\widehat{\pi_{\mathrm{EC}}}$, propose a bootstrap algorithm for it, and show its validity. Exogeneity of budget sets is a standard assumption in classical demand analysis based on random utility models; for example, it is assumed, at least implicitly, in McFadden and Richter (1991). Nonetheless, the assumption can be a concern in applying our testing procedure to a data set such as ours. Recall that the budget sets $\{\mathcal{B}_j\}_{j=1}^J$ are based on prices and total expenditure. The latter is likely to be endogenous, which should be a concern to the econometrician.

As independence between utility and budgets is fundamental to McFadden–Richter theory, addressing it in our testing procedure might seem difficult. Fortunately, recent advances in nonparametric identification and estimation of models with endogeneity inform a solution. To see this, it is useful to rewrite the model so that we can cast it into a framework of nonseparable models with endogenous covariates. Writing $p_j=\tilde{p}_j/W$, where $\tilde{p}_j$ is the unnormalized price vector, the essence of the problem is as follows. Stochastic rationalizability imposes restrictions on the distributions of $y=D(p,u)$ for different p when u is distributed according to its population marginal distribution P_u, but the observed conditional distribution of y given p does not estimate this when w and u are interrelated. In particular, if we define $\mathcal{J}_{\mathrm{EC}}=\min_{\nu\in\mathbf{R}^h_+}[\pi_{\mathrm{EC}}-A\nu]'\Omega[\pi_{\mathrm{EC}}-A\nu]$, with the definition of π_{EC} provided in Section 5, then $\mathcal{J}_{\mathrm{EC}}=0$ iff stochastic rationalizability holds. Note that the new definition π_{EC} recovers the previous definition of π when w is exogenous.

Our estimator uses the control function approach. For example, given a reduced form $w=h_j(z,\varepsilon)$ with h_j monotone in ε and z is an instrument, one may use

$$\varepsilon=F^{(j)}_{w|z}(w|z), \tag{S.8}$$

where $F^{(j)}_{w|z}$ denotes the conditional CDF of w given z under $P^{(j)}$ when the random vector (w,z) obeys the probability law $P^{(j)}$; see Imbens and Newey (2009) for this type of control variable in the context of cross-sectional data. Note that $\varepsilon\sim\mathrm{Uni}(0,1)$ under every $P^{(j)}$, $1\le j\le J$, by construction. Let $P^{(j)}_{y|w,\varepsilon}$ denote the conditional probability measure for y given (w,ε) corresponding to $P^{(j)}$. Adapting the argument in Imbens and Newey (2009) and Blundell and Powell (2003), under the assumption that $\mathrm{supp}(w) = \mathrm{supp}(w|\varepsilon)$ under $P^{(j)}$, $1 \le j \le J$, we have

$$\begin{aligned}\pi(p_j, x_{i|j}) &= \int_0^1 \int_u \mathbf{1}\{D_j(\underline{w}_j, u) \in x_{i|j}\} \, \mathrm{d}P^{(j)}_{u|\varepsilon} \, \mathrm{d}\varepsilon \\ &= \int_0^1 P^{(j)}_{y|w,\varepsilon}\{y \in x_{i|j} | w = \underline{w}_j, \varepsilon\} \, \mathrm{d}\varepsilon, \quad 1 \le j \le J.\end{aligned}$$

This means that π_{EC} can be estimated nonparametrically.

To estimate $\widehat{\pi_{\mathrm{EC}}}$, we can proceed in two steps as follows. The first step is to obtain control variable estimates $\widehat{\epsilon}_{n(j)}$, $n(j) = 1, \dots, N_j$, for each j. For example, let $\widehat{F}^{(j)}_{w|z}$ be a nonparametric estimator for $F_{w|z}$ for a given instrumental variable z in period j. For concreteness, we consider a series estimator as in Imbens and Newey (2002). Let $r^L(z) = (r_{1L}(z), \dots, r_{LL}(z))$, where $r_{\ell L}(z)$, $\ell = 1, \dots, L$ are basis functions. Then define

$$\widehat{F}^{(j)}_{w|z}(w|z) = r^L(z)' \widehat{R}^{-}(j) \sum_{n(j)=1}^{N_j} r^{L(j)}(z_{n(j)}) \mathbf{1}\{w_{n(j)} \le w\} / N_j,$$

where

$$\widehat{R}(j) = \sum_{n(j)=1}^{N_j} r^{L(j)}(z_{n(j)}) r^{L(j)}(z_{n(j)})' / N_j.$$

Let

$$\widetilde{\epsilon}_{n(j)} = \widehat{F}^{(j)}_{w|z}(w_{n(j)}|z_{n(j)}), \quad n(j) = 1, \dots, N_j.$$

Choose a sequence $\upsilon_N \to 0$, $\upsilon_N > 0$, and define $\iota_N(\varepsilon) = (\varepsilon + \upsilon_N)^2 / 4\upsilon_N$. Then let

$$\gamma_N(\varepsilon) = \begin{cases} 1, & \text{if } \varepsilon > 1 + \upsilon_N, \\ 1 - \iota_N(1 - \varepsilon), & \text{if } 1 - \upsilon_N < \varepsilon \le 1 + \upsilon_N, \\ \varepsilon, & \text{if } \upsilon_N \le \varepsilon \le 1 - \upsilon_N, \\ \iota_N(\varepsilon), & \text{if } -\upsilon_N \le \varepsilon \le \upsilon_N, \\ 0, & \text{if } \varepsilon < -\upsilon_N. \end{cases}$$

Then our control variable is $\widehat{\varepsilon}_{n(j)} = \gamma_N(\widetilde{\varepsilon}_{n(j)})$, $n(j) = 1, \dots, N_j$.

The second step is nonparametric estimation of $P^{(j)}_{y|w,\varepsilon}\{y \in x_{i|j} | w = \underline{w}_j, \varepsilon\}$. Let $\widehat{\chi}_{n(j)} = (w_{n(j)}, \widehat{\varepsilon}_{n(j)})'$, $n(j) = 1, \dots, N_j$, for each j. Write $s^{M(j)}(\chi) = (s_{1M(j)}(\chi), \dots, s_{M(j)M(j)}(\chi))'$, where $s_{mM(j)}(\chi)$, $\chi \in \mathbf{R}^{K+1}$, and $m = 1, \dots, M(j)$ are basis functions. Then our estimator for $P^{(j)}_{y|w,\varepsilon}\{y \in x_{i|j} | w = \cdot, \varepsilon = \cdot\}$ evaluated at $\chi = (w, \varepsilon)$ is

$$\begin{aligned}\widehat{P^{(j)}_{y|w,\varepsilon}}\{y \in x_{i|j} | w, \varepsilon\} &= s^{M(j)}(\chi)' \widehat{S}^{-}(j) \sum_{n(j)=1}^{N_j} s^{M(j)}(\widehat{\chi}_{n(j)}) d_{i|j,n(j)} / N_j \\ &= s^{M(j)}(\chi)' \widehat{\alpha}^{M(j)}_i,\end{aligned}$$

where

$$\widehat{S}(j) = \sum_{n(j)=1}^{N_j} s^{M(j)}(\widehat{\chi}_{n(j)}) s^{M(j)}(\widehat{\chi}_{n(j)})'/N_j, \qquad \widehat{\alpha}_i^{M(j)} := \widehat{S}^{-}(j) \sum_{n(j)=1}^{N_j} s^{M(j)}(\widehat{\chi}_{n(j)}) d_{i|j,n(j)}/N_j.$$

Our endogeneity-corrected conditional probability $\pi(p_j, x_{i|j})$ is a linear functional of $P^{(j)}_{y|w,\varepsilon}\{y \in x_{i|j} | w = \underline{w}_j, \varepsilon\}$. Thus, plugging $\widehat{P^{(j)}_{y|w,\varepsilon}}\{y \in x_{i|j} | w = \underline{w}_j, \varepsilon\}$ into the functional, we define

$$\begin{aligned} \pi\widehat{(p_j, x_{i|j})} &:= \int_0^1 \widehat{P^{(j)}_{y|w,\varepsilon}}\{y \in x_{i|j} | w = \underline{w}_j, \varepsilon\} d\varepsilon \\ &= D(j)' \widehat{\alpha}_i^{M(j)}, \end{aligned}$$

where

$$D(j) := \int_0^1 s^{M(j)}\left(\begin{bmatrix} \underline{w}_j \\ \varepsilon \end{bmatrix}\right) d\varepsilon, \quad i = 1, \ldots, I_j, j = 1, \ldots, J,$$

and

$$\begin{aligned} \widehat{\pi_{\mathrm{EC}}} = \big[&\pi\widehat{(p_1, x_{1|1})}, \ldots, \pi\widehat{(p_1, x_{I_1|1})}, \pi\widehat{(p_2, x_{1|2})}, \ldots, \pi\widehat{(p_2, x_{I_2|2})}, \ldots, \\ &\pi\widehat{(p_J, x_{1|J})}, \ldots, \pi\widehat{(p_J, x_{I_J|J})}\big]'. \end{aligned}$$

The final form of the test statistic is

$$\mathcal{J}_{\mathrm{EC}_N} = N \min_{\nu \in \mathbf{R}^h_+} [\widehat{\pi_{\mathrm{EC}}} - A\nu]' \Lambda [\widehat{\pi_{\mathrm{EC}}} - A\nu].$$

The calculation of critical values can be carried out in the same way as the testing procedure with the series estimator $\hat{\pi}$ for the exogenous case, though the covariance matrix $v_N^{(j)}$ needs modification. With the nonparametric endogeneity correction, the modified version of $v_N^{(j)}$ is

$$\bar{v}_N^{(j)} = \big[\mathbf{I}_{I_j} \otimes D(j)' S_N(j)^{-1}\big] \bar{\Lambda}_N^{(j)} \big[\mathbf{I}_{I_j} \otimes S_N(j)^{-1} D(j)\big],$$

where

$$\begin{aligned} S_N(j) &= \mathrm{E}\big[s^{M(j)}(\chi_{n(j)}) s^{M(j)}(\chi_{n(j)})'\big], \qquad \bar{\Lambda}_N^{(j)} = \bar{\Lambda}_{1_N}^{(j)} + \bar{\Lambda}_{2_N}^{(j)}, \\ \bar{\Lambda}_{1_N}^{(j)} &= \mathrm{E}\big[\bar{\Sigma}^{(j)}(\chi_{n(j)}) \otimes s^{M(j)}(\chi_{n(j)}) s^{M(j)}(\chi_{n(j)})'\big], \qquad \bar{\Lambda}_{2_N}^{(j)} = \mathrm{E}\big[m_{n(j)} m'_{n(j)}\big] \end{aligned}$$

with

$$\begin{aligned} \bar{\Sigma}^{(j)}(\chi) &:= \mathrm{Cov}[d_{j,n(j)} | \chi_{n(j)} = \chi], \\ m_{n(j)} &:= \big[m'_{1,n(j)}, m'_{2,n(j)}, \ldots, m'_{I_j,n(j)}\big]', \\ m_{i,n(j)} &:= \mathrm{E}\bigg[\dot{\gamma}_N(\varepsilon_{m(j)}) \frac{\partial}{\partial \varepsilon} P^{(j)}_{y|w,\varepsilon}\{y \in x_{i|j} | w_{m(j)}, \varepsilon_{m(j)}\} \\ &\qquad \times s^{M(j)}(\chi_{m(j)}) r^{L(j)}(z_{m(j)})' R_N(j)^{-1} r^{L(j)}(z_{n(j)}) u_{mn(j)} \bigg| d_{i|j,n(j)}, w_{n(j)}, z_{n(j)}\bigg], \end{aligned}$$

$$R_N(j) := \mathrm{E}\big[r^{L(j)}(z_{n(j)})r^{L(j)}(z_{n(j)})'\big],$$
$$u_{mn(j)} := \mathbf{1}\{w_{n(j)} \le w_{m(j)}\} - F^{(j)}_{w|z}(w_{m(j)}|z_{n(j)}).$$

Define

$$\underline{n}_{\mathrm{EC}} = \min_j N_j I_j / \operatorname{trace}\big(\bar{v}^{(j)}_N\big).$$

Then a possible choice for τ_N is $\tau_N = \sqrt{\frac{\log \underline{n}_{\mathrm{EC}}}{\underline{n}_{\mathrm{EC}}}}$. Proceed as for $\tilde{\mathcal{J}}_N$ earlier in this section, replacing $\hat{v}^{(j)}_N$ with a consistent estimator for $\bar{v}^{(j)}_N$, for $j = 1, \dots, J$, to define the bootstrap version $\tilde{\mathcal{J}}_{\mathrm{EC}}$.

We impose some conditions to show the validity of the endogeneity-corrected test. Let $\varepsilon_{n(j)}$ be the value of the control variable ε for the $n(j)$th consumer facing budget j. Noting that $\{\{d_{i|j,n(j)}\}_{i=1}^{I_j}, w_{n(j)}, \varepsilon_{n(j)}\}_{n(j)=1}^{N_j}$ are IID distributed within each time period j, $1 \le j \le J$, let $(d_j, w_j, \varepsilon_j)$ denote the choice–log-expenditure–control variable triplet of a consumer facing budget j. Let $d = [d_1', \dots, d_J']'$, $\mathbf{w} = [w_1, \dots, w_J]'$, and $\mathbf{e} = [\varepsilon_1, \dots, \varepsilon_J]'$, and define $g = Bd = [g_1, \dots, g_m]'$ as before. Let $\mathcal{X}_j = \operatorname{supp}(\chi_{n(j)})$, $\mathcal{Z}_j = \operatorname{supp}(z_{n(j)})$, and $\mathcal{E}_j = \operatorname{supp}(\varepsilon_{n(j)})$, $1 \le j \le J$. Following the above discussion, define an $\mathbf{R}^I$-valued functional

$$\pi\big(P^{(1)}_{y|w,\varepsilon}, \dots, P^{(J)}_{y|w,\varepsilon}\big) = \big[\pi_{1|1}\big(P^{(1)}_{y|w,\varepsilon}\big), \dots, \pi_{I_1|1}\big(P^{(1)}_{y|w,\varepsilon}\big), \pi_{1|2}\big(P^{(2)}_{y|w,\varepsilon}\big), \dots, \pi_{I_2|2}\big(P^{(2)}_{y|w,\varepsilon}\big), \dots,$$
$$\pi_{1|J}\big(P^{(J)}_{y|w,\varepsilon}\big), \dots, \pi_{I_J|J}\big(P^{(J)}_{y|w,\varepsilon}\big)\big]',$$

where

$$\pi_{i|j}\big(P^{(j)}_{y|w,\varepsilon}\big) := \int_0^1 P^{(j)}_{y|w,\varepsilon}\{y \in x_{i|j} | w = \underline{w}_j, \varepsilon\}\, \mathrm{d}\varepsilon$$

and $\varepsilon_{n(j)} := F^{(j)}_{w|z}(w_{n(j)}|z_{n(j)})$ for every j.

CONDITION S.4: There exist positive constants c_1, c_2, δ_1, δ, $\zeta_r(L)$, $\zeta_s(M)$, and $\zeta_1(M)$, $L \in \mathbf{N}$, $M \in \mathbf{N}$ such that the following statements hold:

(i) The distribution of $w_{n(j)}$ conditional on $z_{n(j)} = z$ is continuous for every $z \in \mathcal{Z}_j$, $1 \le j \le J$.

(ii) We have $\operatorname{supp}(w_{n(j)}|\varepsilon_{n(j)} = \varepsilon) = \operatorname{supp}(w_{n(j)})$ for every $\varepsilon \in [0, 1]$, $1 \le j \le J$.

(iii) We have $\pi(P^{(1)}_{y|w,\varepsilon}, \dots, P^{(J)}_{y|w,\varepsilon}) \in \mathcal{C}$.

(iv) For each $k \in \mathcal{K}^R$, $\operatorname{var}(g_k|\mathbf{w} = (\underline{w_1}, \dots, \underline{w_J})', \mathbf{e} = (\underline{\varepsilon_1}, \dots, \underline{\varepsilon_J}))' \ge s^2(F_1, \dots, F_J)$ and $\mathrm{E}[(g_k/s(F_1, \dots, F_J))^4|\mathbf{w} = (\underline{w_1}, \dots, \underline{w_J})', \mathbf{e} = (\underline{\varepsilon_1}, \dots, \underline{\varepsilon_J})'] < c_1$ hold for every $(\underline{w_1}, \dots, \underline{w_J}, \underline{\varepsilon_1}, \dots, \underline{\varepsilon_J}) \in \mathcal{W}_1 \times \dots \times \mathcal{W}_J \times \mathcal{E}_1 \times \dots \times \mathcal{E}_J$.

(v) Letting $\widetilde{r}^L := C_{L,j} r^L$, $\lambda_{\min} \mathrm{E}[\widetilde{r}^L(z_{n(j)})\widetilde{r}^L(z_{n(j)})] \ge c_2$ holds for every L and j, where $C_{L,j}$, $L \in \mathbf{N}$, $1 \le j \le J$, are constant nonsingular matrices.

(vi) We have $\max_j \sup_{z \in \mathcal{Z}_j} \|\widetilde{r}^L(z)\| \le c_1 \zeta_r(L)$ for every $L \in \mathbf{N}$.

(vii) We have that $\sup_{w \in \mathcal{W}_j, z \in \mathcal{Z}_j} |F^{(j)}_{w|z}(w, z) - r^L(z)'\alpha^{(j)}_L(w)| \le c_1 L^{-\delta_1}$, $1 \le j \le J$ holds with some L vector $\alpha^{(j)}_L(\cdot)$ for every $L \in \mathbf{N}$, $1 \le j \le J$.

(viii) Letting $\widetilde{s}^M := \bar{C}_{M,j} s^M$, $\lambda_{\min} \mathrm{E}[\widetilde{s}^M(\chi_{n(j)})\widetilde{s}^M(\chi_{n(j)})] \ge c_2$ holds for every M and j, where $\bar{C}_{M,j}$, $M \in \mathbf{N}$, $1 \le j \le J$, are constant nonsingular matrices.

(ix) We have $\max_j \sup_{\chi \in \mathcal{X}_j} \|\widetilde{s}^M(\chi)\| \le C\zeta_s(M)$ and $\max_j \sup_{\chi \in \mathcal{X}_j} \|\partial \widetilde{s}^M(\chi)/\partial \varepsilon\| \le c_1 \zeta_1(M)$, and $\zeta_s(M) \le C\zeta_1(M)$ for every $M \in \mathbf{N}$.

(x) We have that $\sup_{\chi\in\mathcal{X}_j}|P^{(j)}_{y|w,\varepsilon}\{y\in x_{i|j}|w,\varepsilon\}-s^M(\chi)'g^{(i,j)}_M|\le c_1M^{-\delta}$ holds with some M vector $g^{(i,j)}_M$ for every $M\in\mathbf{N}$, $1\le i\le I_j$, $1\le j\le J$.

(xi) The $P^{(j)}_{y|w,\varepsilon}\{y\in x_{i|j}|w,\varepsilon\}$, $1\le i\le I_j$, $1\le j\le J$, are twice continuously differentiable in $\chi=(w,\varepsilon)$. Moreover, $\max_{1\le j\le J}\max_{1\le i\le I_j}\sup_{\chi\in\mathcal{X}_j}\|\frac{\partial}{\partial\chi}P^{(j)}_{y|w,\varepsilon}\{y\in x_{i|j}|w,\varepsilon\}\|\le c_1$ and $\max_{1\le j\le J}\max_{1\le i\le I_j}\sup_{\chi\in\mathcal{X}_j}\|\frac{\partial^2}{\partial\chi\partial\chi'}P^{(j)}_{y|w,\varepsilon}\{y\in x_{i|j}|w,\varepsilon\}\|\le c_1$.

Since we use the control function approach to deal with potential endogeneity in w (income), Condition S.4(i) and (ii) are essential. See Blundell and Powell (2003) and Imbens and Newey (2009) for further discussion on these types of restrictions. Just like Condition S.2(ii), Condition S.4(iv) is a version of Condition S.1 that accommodates the two-step series estimation. Conditions S.4(iv)–(xi) correspond to standard regularity conditions stated in the context of the two-step approach adopted in this section: (iv) and (x) imposes uniform approximation error bounds; (v) and (viii) avoid singular designs (note the existence of the matrices $C_{L,j}$ and $\bar{C}_{M,j}$ suffices); (vi) and (ix) control the lengths of (the derivatives of) the series terms used. Condition S.4(xi) imposes reasonable smoothness restrictions on the (observable) conditional probabilities $P^{(j)}_{y|w,\varepsilon}\{y\in x_{i|j}|w,\varepsilon\}$, $1\le i\le I_j$, $1\le j\le J$.

The next condition impose restrictions on tuning parameters.

CONDITION S.5: Let τ_N, $M(j)$, and $L(j)$, $j=1,\dots,J$, satisfy $\tau_N\downarrow 0$, $\sqrt{\underline{n}_{\mathrm{EC}}}\tau_N\uparrow\infty$, $N_jL(j)^{1-2\delta_1}\downarrow 0$, $N_jM(j)^{-2\delta}\downarrow 0$, $M(j)\zeta_1(M(j))^2L^2(j)/N_j\downarrow 0$, $\zeta_s(M(j))^6L^4(j)/N_j\downarrow 0$, and $\zeta_1(M(j))^4\zeta_r(L(j))^4/N_j\downarrow 0$, and also $\underline{C}(L(j)/N_j+L(j)^{1-2\delta_1})\le v_N^3\le\overline{C}(L(j)/N_j+L(j)^{1-2\delta_1})$, for some $0<\underline{C}<\overline{C}$.

PROOF OF THEOREM 5.2: The proof follows the same steps as those in the proof of Theorem 4.2, except for the treatment of the estimator for π. Therefore, instead of the sequence π_N, $N=1,2,\dots$, in $\mathcal{P}\cap\mathcal{C}$, consider a sequence of distributions $F_N=[F_{1N},\dots,F_{JN}]$, $N=1,2,\dots$, in $\mathcal{F}_{\mathrm{EC}}$ and the corresponding conditional distributions $P^{(j)}_{y|w,\varepsilon;F_N}\{y\in x_{i|j}|w,\varepsilon\}$ and $F^{(j)}_{w|z_N}$, $1\le i\le I_j$, $1\le j\le J$, $N=1,2,\dots$, such that $\sqrt{N_j/(M(j)\vee L(j))}B^{(j)}\pi^{(j)}_{F_N}\to h_j$, $h_j\le 0$, $1\le j\le J$ as $N\to\infty$, where $\pi_{F_N}=\pi(P^{(1)}_{y|w,\varepsilon_N},\dots,P^{(J)}_{y|w,\varepsilon_N})$, whereas the definitions of $\overline{V}^{(j)}_{F_N}$, $1\le j\le J$, are given shortly. Define $S^{(j)}_{F_N}=\mathrm{E}_{F_N}[s^{M(j)}(\chi_{n(j)})s^{M(j)}(\chi_{n(j)})']$ as well as

$$\bar{\Xi}^{(j)}_{1_{F_N}}=\mathrm{E}_{F_N}\big[B^{(j)}\bar{\Sigma}^{(j)}_{F_N}(\chi_{n(j)})B^{(j)\prime}\otimes s^{M(j)}(\chi_{n(j)})s^{M(j)}(\chi_{n(j)})'\big]$$

and

$$\bar{\Xi}^{(j)}_{2_{F_N}}=\big[B^{(j)}\otimes\mathbf{I}_{M(j)}\big]\mathrm{E}_{F_N}\big[m_{n(j);F_N}m'_{n(j);F_N}\big]\big[B^{(j)\prime}\otimes\mathbf{I}_{M(j)}\big],$$

where

$$\Sigma^{(j)}_{F_N}(\chi):=\mathrm{Cov}_{F_N}[d_{j,n(j)}|\chi_{n(j)}=\chi],$$

$$m_{n(j);F_N}:=\big[m'_{1,n(j);F_N},m'_{2,n(j);F_N},\dots,m'_{I_j,n(j);F_N}\big]',$$

$$m_{i,n(j);F_N}:=\mathrm{E}_{F_N}\bigg[\dot{\gamma}_N(\varepsilon_{m(j)})\frac{\partial}{\partial\varepsilon}P^{(j)}_{y|w,\varepsilon;F_N}\{y\in x_{i|j}|w_{m(j)},\varepsilon_{m(j)}\}$$

$$\times s^{M(j)}(\chi_{m(j)}) r^{L(j)}(z_{m(j)})' R_{F_N}(j)^{-1} r^{L(j)}(z_{n(j)}) u_{mn(j);F_N} \bigg| d_{i|j,n(j)}, w_{n(j)}, z_{n(j)} \bigg],$$

$$R_{F_N}(j) := \mathrm{E}_{F_N}\big[r^{L(j)}(z_{n(j)}) r^{L(j)}(z_{n(j)})'\big],$$

$$u_{mn(j);F_N} := \mathbf{1}\{w_{n(j)} \leq w_{m(j)}\} - F^{(j)}_{w|z_N}(w_{m(j)}|z_{n(j)}).$$

With these definitions, let

$$\overline{V}^{(j)}_{F_N} := \big[\mathbf{I}_m \otimes D(j)' S^{(j)\,-1}_{F_N}\big] \bar{\Xi}^{(j)}_{F_N} \big[\mathbf{I}_m \otimes S^{(j)\,-1}_{F_N} D(j)\big]$$

with $\bar{\Xi}^{(j)}_{F_N} = \bar{\Xi}^{(j)}_{1_{F_N}} + \bar{\Xi}^{(j)}_{2_{F_N}}$. Define

$$\overline{V}_{F_N} := \sum_{j=1}^{J} \overline{V}^{(j)}_{F_N}.$$

Then by adapting the proof of Theorem 7 in Imbens and Newey (2002) to the triangular array for the repeated cross section setting, for the js that satisfy Condition (iv), we obtain

$$\sqrt{N}\overline{V}_{F_N}^{-\frac{1}{2}} B[\hat{\pi} - \pi_{F_N}] \overset{F_N}{\rightsquigarrow} N(0, \mathbf{I}_m).$$

The rest is the same as the proof of Theorem 4.2. *Q.E.D.*

APPENDIX B: ALGORITHMS FOR COMPUTING A

This appendix details algorithms for computation of A. The first algorithm is the depth-first search that we in fact implemented. The second algorithm is a further refinement using Theorem 3.2. Algorithms use notation introduced in the proof of Theorem 3.2.

Computing A as in Theorem 3.1

1. Initialize $m_1 = \cdots = m_J = 1$.
2. Initialize $l = 2$.
3. Set $c(\mathcal{B}_1) = x_{m_1|1}, \ldots, c(\mathcal{B}_l) = x_{m_l|l}$. Check for revealed preference cycles.
4. If a cycle is detected, move to step 7. Else:
5. If $l < J$, set $l = l + 1$, $m_l = 1$, and return to step 3. Else:
6. Extend A by the column $[m_1, \ldots, m_J]'$.

7a. If $m_l < I_l$, set $m_l = m_l + 1$ and return to step 3.

7b. If $m_l = I_l$ and $m_{l-1} < I_{l-1}$, set $m_l = 1$, $m_{l-1} = m_{l-1} + 1$, $l = l - 1$, and return to step 3.

7c. If $m_l = I_l$, $m_{l-1} = I_{l-1}$, and $m_{l-2} < I_{l-2}$, set $m_l = m_{l-1} = 1$, $m_{l-2} = m_{l-2} + 1$, $l = l - 2$, and return to step 3.

(...)

7z. Terminate.

Refinement Using Theorem 3.2

Let budgets be arranged such that $(\mathcal{B}_1, \ldots, \mathcal{B}_M)$ do not intersect $\mathcal{B}_J$; for exposition of the algorithm, assume $\mathcal{B}_J$ is above these budgets.

```
1. Use preceding algorithm to compute a matrix A_{M+1→J-1} corre-
sponding to budgets (B_{M+1},...,B_J), though using the full X
corresponding to budgets (B_1,...,B_J).
2. For each column a_{M+1→J-1} of A_{M+1→J-1}, go through the follow-
ing steps:
2.1 Compute (using preceding algorithm) all vectors a_{1→M} s.t.
(a_{1→M}, a_{M+1→J-1}) is rationalizable.
2.2 Compute (using preceding algorithm) all vectors a_J s.t.
(a_{M+1→J-1}, a_J) is rationalizable.
2.3 All stacked vectors (a'_{1→M}, a'_{M+1→J-1}, a'_J)' are valid columns
of A.
```
[3]

APPENDIX C: JUSTIFICATION OF TABLE I

This appendix derives the upper bound on nodes visited by a tree search as described in Section 3.4.2. We only count nodes corresponding to $j \geq 2$, as rationalizability of implied choice behavior is checked only at those.

Consider the number of nodes visited in generation $j+1$, that is, corresponding to budget $\mathcal{B}_j$. Since $I_j \leq 2^{J-1}$, this is at most 2^{J-1} times the number of nodes in the jth generation at which no choice cycle was detected. These nodes, in turn, correspond to the at most $\bar{H}_{j-1}$ direct revealed preference orderings that can occur on $(j-1)$ budgets. However, since we look at patches corresponding to the entire set of J budgets, each of those orderings has multiple representations. Specifically, each patch in an A matrix corresponding to the first $(j-1)$ budgets corresponds to at most $2^{J-(j-1)}$ patches in the problem under consideration (because the patches are generated by intersecting the original patch with $(J-(j-1))$ budgets). These refined patches can be arbitrarily combined across the first $(j-1)$ budgets, so that each direct revealed preference ordering on the first $(j-1)$ budgets has at most $2^{(J-j+1)(j-1)}$ representations. Thus, the number of nodes visited in generation $j+1, j, \dots, J$, is at most $\bar{H}_{j-1}2^{(J-j+1)(j-1)+J-1} = \bar{H}_{j-1}2^{j(J+2-j)-2}$. This bound must be summed over $j = 2, \dots, J$.

REFERENCES

ANDREWS, D. W., AND G. SOARES (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157.[4,9]

BLUNDELL, R., AND J. L. POWELL (2003): "Endogeneity in Nonparametric and Semiparametric Regression Models," *Econometric Society Monographs*, 36, 312–357.[13,16]

IMBENS, G. W., AND W. K. NEWEY (2002): "Identification and Estimation of Triangular Simultaneous Equation Models Without Additivity," NBER Technical Working Paper Series. [13,17]

——— (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77 (5), 1481–1512.[10,12,16]

MCFADDEN, D., AND K. RICHTER (1991): "Stochastic Rationality and Revealed Stochastic Preference," in *Preferences, Uncertainty and Rationality*, ed. by J. Chipman, D. McFadden, and K. Richter. Boulder: Westview Press, 161–186. [12]

NEWEY, W. K. (1997): "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, 79 (1), 147–168.[9,12]

Co-editor Elie Tamer handled this manuscript.

Manuscript received 15 June, 2016; final version accepted 30 May, 2018; available online 25 July, 2018.

[3]This matrix has more rows than an A matrix that is only intended to apply to choice problems $(\mathcal{B}_{M+1}, \dots, \mathcal{B}_J)$.