Supplemental Appendix:
Extensions, Variations, and Robustness

In this Appendix we explore variations on our baseline model and associated identification conditions. We show that identification is robust in the sense that a relaxation of one condition assumed in the text can often be accommodated by strengthening another. An understanding of such trade-offs can be helpful to both producers and consumers of research relying on demand estimates. Although a full exploration of these trade-offs describes an entire research agenda, we illustrate some possibilities that relax key restrictions of our model, allow demand systems outside discrete choice settings, enlarge the set of potential instruments, reduce the number of required instruments, eliminate the need for continuous consumer-level observables, or reduce the required dimension of those observables. For simplicity, we focus here on the traditional case in which X_t is exogenous, recalling that in this case we have $h(X_t, \Xi_t) = \Xi_t$.

S.1 Prices in the Index

In the text we excluded prices from the index vector $\gamma\left(Z_{it}, X_t, \Xi_t\right)$. That did not rule out interactions between prices and individual-specific measures (e.g., income), but required that such measures be "extra" consumer-level observables beyond those in the J-vector Z_{it}. That requirement can be relaxed. A full investigation of identification in models allowing interactions between Z_{it} and P_t is beyond the scope of this project. However, here we discuss one class of fully nonparametric models permitting such interactions, demonstrating one direction in which our results can be extended.

S.1.1 Model and Normalizations

Suppose demand takes the form

$$\sigma\left(\gamma\left(Z_{it}, P_t, X_t, \Xi_t\right), P_t, X_t\right), \tag{S.3}$$

where, for each $j = 1, \ldots, J$,

$$\gamma_j\left(Z_{it}, P_t, X_t, \Xi_t\right) = g_j\left(Z_{it}, P_{jt}, X_t\right) + \Xi_{jt},$$

$$g_j\left(Z_{it}, P_{jt}, X_t\right) = \bar{g}_j\left(Z_{it}, X_t\right) + \tilde{g}_j(\tilde{Z}_{it}, P_{jt}, X_t), \tag{S.4}$$

and

$$\tilde{Z}_{it} \equiv (Z_{i2t}, \ldots, Z_{iJt}).$$

This specification imposes two restrictions on an otherwise fully-flexible index function $\gamma(Z_{it}, P_t, X_t, \Xi_t)$.[1] First, the price of good j affects only the index $\gamma_j(Z_{it}, P_t, X_t, \Xi_t)$ associated with good j. This is a natural restriction implied by standard specifications (see, e.g., section 3). Second, at least one element of Z_{it} is excluded from interacting with prices. This is an important restriction but also standard in practice. The key implication, exploited below, is that $\frac{\partial g(z,p)}{\partial z_1}$ does not vary with p.

Given the conditions in the text, identification in this model can be obtained under additional verifiable conditions. Here we sketch the argument. With exogenous X_t we can condition on X_t and drop it from the notation.[2]

This model requires a slightly different set of normalizations from those used in the text. Let z^0 denote an arbitrary point in $\mathcal{Z}$ for which $\partial s_t(z^0)/\partial z$ is nonsingular, and let $\tilde{z}^0 = (z_2^0, \ldots, z_J^0)$. Without loss of generality, for each $j = 1, \ldots, j$ we set

$$E[\Xi_{jt}] = 0 \tag{S.5}$$

$$\tilde{g}_j\left(\tilde{z}^0, p_{jt}\right) = 0 \quad \forall p_{jt} \tag{S.6}$$

$$\bar{g}_j\left(z^0\right) = 0 \tag{S.7}$$

$$\frac{\partial \bar{g}_j\left(z^0\right)}{\partial z_1} = 1. \tag{S.8}$$

Equation (S.5) normalizes the location of Ξ_t, as in the text. The need for (S.6) reflects the fact that each price P_{jt} already appears in unrestricted form in the function σ. The role of prices in the index vector is to allow variation in price responses across consumers with different values of $\tilde{Z}_{it}$. Equation (S.6) defines $\tilde{z}^0$ as the (arbitrary) baseline value of $\tilde{Z}_{it}$ around which such variation is defined.[3] Given (S.5) and (S.6), (S.7) normalizes the location of each index γ_j, while (S.8) normalizes its scale (see the related discussion in the text).

---

[1] With no restriction on $\gamma(Z_{it}, P_t, X_t, \Xi_t)$, (S.3) would impose no restriction on demand as function of $(Z_{it}, P_t, X_t, \Xi_t)$, and there would be no role for the function σ in (S.3).

[2] Conditioning on X_t treats it fully flexibly, as the argument presented can be applied at each value of X_t.

[3] More formally, for any demand system $\sigma(\gamma(Z_{it}, P_t, \Xi_t), P_t)$ and any functions $\kappa_j(P_{jt})$ for each j, one can can define an equivalent representation of this demand as $\hat{\sigma}(\hat{\gamma}(Z_{it}, P_t, \Xi_t), P_t)$, where $\hat{\gamma}_j(Z_{it}, P_{jt}, \Xi_{jt}) \equiv \gamma_j(Z_{it}, P_{jt}, \Xi_{jt}) - \kappa_j(P_{jt})$ and $\hat{\sigma}(\hat{\gamma}, P_t) \equiv \sigma(\hat{\gamma} + \kappa(P_t), P_t)$. We select the representation with $\kappa_j(P_{jt}) = \tilde{g}_j\left(\tilde{z}_{it}^0, P_{jt}\right)$ for all j.

S.1.2 Identification: Sketch

Following the arguments in Lemmas 1–3 in the text, one can show that for every price p and any $z' \in \mathcal{Z}$ the value of

$$V\left(z', z^0, p\right) \equiv \left(\frac{\partial g(z^0, p)}{\partial z}\right)^{-1} \left(\frac{\partial g(z', p)}{\partial z}\right) \tag{S.9}$$

is identified, although the two matrices on the RHS are unknown. Thus, (S.9) provides a system of J^2 equations in the $2J^2$ elements of these matrices. Rewrite this system as

$$\frac{\partial g(z^0, p)}{\partial z} V(z', z^0, p) = \frac{\partial g(z', p)}{\partial z}. \tag{S.10}$$

These J^2 equations break naturally into J groups, each with form

$$\underbrace{\frac{\partial g_j(z^0, p)}{\partial z}}_{1\times J} \underbrace{V(z', z^0, p)}_{J\times J} = \underbrace{\frac{\partial g_j(z', p)}{\partial z}}_{1\times J}. \tag{S.11}$$

Take any $j \in \{1, \ldots, J\}$. A key observation is that one obtains a new system of J equations of the form (S.11) at every new price vector. By choosing prices carefully, this can provide new equations without new unknowns. Starting from an arbitrary price vector p and (S.11), any price vector $p^j \neq p$ for which $p^j_j = p_j$ yields

$$\frac{\partial g_j(z^0, p^j)}{\partial z} V(z', z^0, p^j) = \frac{\partial g_j(z', p^j)}{\partial z}. \tag{S.12}$$

Because g_j depends on the price vector only through p_j, we have $\frac{\partial g_j(z^0,p^j)}{\partial z} = \frac{\partial g_j(z^0,p)}{\partial z}$ and $\frac{\partial g_j(z',p^j)}{\partial z} = \frac{\partial g_j(z',p)}{\partial z}$. So we may rewrite (S.12) as

$$\frac{\partial g_j(z^0, p)}{\partial z} V(z', z^0, p^j) = \frac{\partial g_j(z', p)}{\partial z}. \tag{S.13}$$

Subtracting (S.13) from (S.11), we obtain

$$\frac{\partial g_j(z^0, p)}{\partial z} \Lambda\left(z', z^0, p, p^j\right) = 0 \tag{S.14}$$

where

$$\Lambda\left(z', z^0, p, p^j\right) \equiv V(z', z^0, p) - V(z', z^0, p^j). \tag{S.15}$$

Equation (S.14) is a homogeneous system of J linear equations in the J

components of $\frac{\partial g_j(z^0,p)}{\partial z}$. One of these elements is already known: by (S.4) and (S.8), $\frac{\partial g_j(z^0,p)}{\partial z_1} = 1$. One can solve (S.14) for the remaining elements under the following condition on $\Lambda(z', z^0, p, p^j)$.

Condition 1. *For some $k \in \{1, \ldots, J\}$, the submatrix of $\Lambda(z', z^0, p, p^j)$ obtained by dropping row j and column k is full rank.*

Example 1. *Consider the case of $J = 2$ and $j = 1$. Letting $d_{jk} = \partial g_j(z^0, p)/\partial z_k$, (S.14) takes the form (recalling the normalization $d_{11} = 1$)*

$$\begin{bmatrix} 1 & d_{12} \end{bmatrix} \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} = 0,$$

which we may rewrite as

$$\Lambda_{11} + d_{12}\Lambda_{21} = 0$$
$$\Lambda_{12} + d_{12}\Lambda_{22} = 0.$$

Either of these equations could be used to solve for the unknown d_{12}, although this requires that at least one of Λ_{12} and Λ_{22} be nonzero (Condition 1). If $\Lambda_{22} \neq 0$, for example, then $d_{12} = -\frac{\Lambda_{12}}{\Lambda_{22}}$.

Repeating the argument above for each j identifies the matrix $\frac{\partial g(z^0,p)}{\partial z}$. Plugging this into (S.10) (and recalling that $V(z', z^0, p)$ is known) then allows identification of $\frac{\partial g(z',p)}{\partial z}$ at every z'. Since the normalizations (S.6) and (S.7) imply the boundary condition $g(z^0, p) = 0$, we can integrate $\frac{\partial g(z,p)}{\partial z}$ from this point to identify $g(z, p)$ at all z. Repeating the entire argument at each p then identifies the function g. With g known, the arguments in the text (starting from Corollary 1) can be applied directly to show identification of demand.

Thus, our identification results extend when, in addition to the conditions in the text, for every price vector p and each $j = 1, \ldots, J$, there exists a pair (z', p^j) with $p^j_j = p_j$ and for which Condition 1 is satisfied. Because P_t is observed and Condition 1 is a property of identified objects, this requirement is verifiable.[4] Condition 1 requires that the price vector alter the derivative matrix $\frac{\partial g(z,p)}{\partial z}$, and differentially so at different prices p. This requires nonlinearity: it will fail, for example if g is everywhere linear in z at each p. However, one can confirm numerically that that Condition 1 holds in nonlinear examples—e.g., in specifications following Berry, Levinsohn, and Pakes (1995, 2004). An interesting question is whether there are useful sufficient conditions for Condition 1. That is a topic we leave to future work.

[4] See Berry and Haile (2018) for a formal definition of verifiability.

S.2 Strengthening the Index Structure

The model used by Berry and Haile (2014) to study identification with market-level data restricted the way some elements of X_t enter. Partitioning X_t as $(X_t^{(1)}, X_t^{(2)})$, where $X_t^{(1)} = \left(X_{1t}^{(1)}, \ldots, X_{Jt}^{(1)}\right) \in \mathbb{R}^J$, they assumed that each $X_{jt}^{(1)}$ affects demand only through the jth element of the index vector. This structure is common in specifications used in practice, and adding it here can allow the use of BLP instruments for prices.[5]

To illustrate, suppose demand takes the form

$$\mathscr{s}\left(Z_{it}, P_t, X_t, \Xi_t\right) = \sigma\left(\gamma\left(Z_{it}, X_t, \Xi_t\right), P_t, X_t^{(2)}\right), \tag{S.16}$$

where for each $j = 1, \ldots, J$

$$\gamma_j\left(Z_{it}, X_t, \Xi_t\right) = g_j(Z_{ijt}, X_t^{(2)}) + \eta_j(X_{jt}^{(1)}, X_t^{(2)}) + \Xi_{jt}, \tag{S.17}$$

with $\partial g_j(z, x^{(2)})/\partial z_j > 0$ for all $(z, x^{(2)})$. Compared to the model in the text, this introduces the exclusivity restriction on each $X_{jt}^{(1)}$, associates each Z_{ijt} exclusively with the jth element of the index as well, and imposes separability between Z_{ijt} and $X_{jt}^{(1)}$ within the index.[6] Many specifications in the literature satisfy these requirements, typically with additional restrictions such as linear substitution between Z_{ijt} and $X_{jt}^{(1)}$.

For the remainder of this section we condition on $X_t^{(2)}$ (treating it fully flexibly), suppress it from the notation, and let X_t represent $X_t^{(1)}$. For each $p \in \operatorname{supp} P_t$, define

$$\begin{aligned}
\mathcal{S}(p) &= \bigcup_{x \in \operatorname{supp} X_t | \{P_t = p\}} \mathcal{S}(p, x) \\
\underline{\mathcal{S}}(p) &= \bigcap_{x \in \operatorname{supp} X_t | \{P_t = p\}} \mathcal{S}(p, x) \\
\underline{\mathcal{Z}} &= \bigcap_{x \in \mathcal{X}} \mathcal{Z}(x).
\end{aligned}$$

We assume that $\underline{\mathcal{Z}}$ is nonempty, as is $\underline{\mathcal{S}}(p)$ for all $p \in \operatorname{supp} P_t$. Nonempty $\underline{\mathcal{S}}(p)$ requires that there exist $\underline{s}(p) \in \mathcal{S}(p)$ such that at each $x \in \operatorname{supp} X_t | \{P_t = p\}$

[5] As suggested in section 5.3, the key issue is proper excludability of these instruments, not their relevance.

[6] Exclusivity of $X_{jt}^{(1)}$ to the index γ_j is essential to the point we illustrate here, and this is most natural when exclusivity of each Z_{ijt} differentiates the elements of the index vector. The assumed separability simplifies the analysis.

there is a combination of Z_{it} and Ξ_t in their support conditional on $\{P_t = p, X_t = x\}$ that will map to the choice probability $\underline{s}(p)$. Nonempty $\underline{\mathcal{Z}}$ requires that there exist at least one value of Z_{it} that is present in all markets.

With this more restrictive model we must revisit the necessary normalizations. First, because adding a constant κ_j to g_j and subtracting the same constant from η_j would leave the demand function unchanged, we take an arbitrary $x^0 \in \operatorname{supp} X_t$ and set

$$\eta_j(x_j^0) = 0 \quad \forall j \tag{S.18}$$

without loss. Even with (S.18) (and our maintained $\mathbb{E}[\Xi_t] = 0$), it remains true that linear transformations of each index function γ_j could be offset by an appropriate adjustment to the function σ, yielding multiple representations of the same demand system (recall the related observation in section 2.5). Thus, without loss, we normalize the location and scale of each index by taking an arbitrary $z^0 \in \underline{\mathcal{Z}}$ and setting $g_j\left(z_j^0\right) = 0$ and $\frac{\partial g_j\left(z_j^0\right)}{\partial z_j} = 1$ for all j.

The arguments in Lemmas 1–3 will now demonstrate identification of each function g_j. Likewise, for each price vector p and arbitrary s^0 and s^1 in $\mathcal{S}(p)$, the arguments in Corollary 1 imply identification of

$$\Omega(s^1, s^0, p) \equiv \sigma^{-1}(s^1; p) - \sigma^{-1}(s^0; p).$$

Taking an arbitrary $\tilde{z}_{it} \in \mathcal{Z}(x_t)$ for each market t, the inverted demand system (cf. equation (19)) in each market takes the form

$$g_j\left(\tilde{z}_{ijt}\right) + \eta_j\left(x_{jt}\right) + \xi_{jt} = \sigma_j^{-1}\left(s_t(\tilde{z}_{it}); p_t\right) \quad j = 1, \ldots, J.$$

Taking an arbitrary $s^0(p) \in \mathcal{S}(p)$ at each price vector p, we can rewrite the jth equation as

$$g_j\left(\tilde{z}_{ijt}\right) - \Omega(s_t(\tilde{z}_{it}), s^0(p_t), p_t) = -\eta_j\left(x_{jt}\right) + \sigma_j^{-1}\left(s^0(p_t); p_t\right) - \xi_{jt}. \tag{S.19}$$

Because the LHS is known, this takes the form of a nonparametric regression equation with RHS variables x_{jt} and p_t. In this equation x_{-jt} is excluded, offering $J-1$ potential instruments for the endogenous prices p_t. Thus, one additional instrument—e.g., a scalar market-level cost shifter or Waldfogel instrument—could yield enough instruments to obtain identification of the unknown RHS functions and the "residuals" ξ_{jt}. Once these demand shocks are identified, identification of demand follows immediately.

Many variations on this structure are possible. For example, as in many empirical specifications, one might assume that p_{jt} enters demand only through the j^{th} index. Strengthening the assumption of nonempty $\underline{\mathcal{S}}(p)$ to require nonempty $\bigcap_{(p,x)\in\mathcal{PX}} S(p,x)$, this can lead to a regression equation (the analog of (S.19)) of the form

$$g_j\left(\tilde{z}_{ijt}\right) - \Omega(s_t(\tilde{z}_{it}), s^0) = -\eta_j\left(x_{jt}, p_{jt}\right) + \sigma_j^{-1}\left(s^0\right) - \xi_{jt},$$

where s^0 is an arbitrary point in $\bigcap_{(p,x)\in\mathcal{PX}} S(p,x)$ and the LHS is known. Now only one instrument for price is necessary. For example, the BLP instruments can overidentify demand.

S.3 A Nonparametric Special Regressor

A different approach is to assume that the demand system of interest is generated by a random utility model with conditional indirect utilities of the form

$$U_{ijt} = g_j(Z_{ijt}) + \Xi_{jt} + \mu_{ijt}, \tag{S.20}$$

where μ_{ijt} is a scalar random variable whose nonparametric distribution depends on X_{jt} and P_{jt} (equation (16) gives a parametric example). In this case, our Lemma 3 demonstrates identification of each function $g_j(\cdot)$ up to a normalization of utilities. Under the assumptions of Theorems 1 and 2, conditional demand and demand are identified as the main body of the text.

In the special case of equation (S.20), there is an alternate route to identification. Adding the assumption of independence between the vector μ_{it} and (Ξ_t, Z_{it}) then turns each $g_j(Z_{ijt})$ into a known special regressor. Under a further (and typically very strong) large support assumption on $g(Z_{it})$, standard arguments lead to identification of the marginal distribution of $(\xi_{jt} + \mu_{ijt})|(X_t, P_t)$ for each j in each market t. One can then use these marginal distributions to define a cross-market nonparametric IV regression equation for each choice j, where the LHS is a conditional mean and Ξ_{jt} appears on the RHS as an additive structural error.[7] In each of these equations the prices and characteristics of goods $k \neq j$ are excluded. Identification of the regression functions identifies all demand shocks, and identification of demand then follows as in Theorem 2. Thus, here one needs only one instrument for price, and exogenous characteristics of competing goods (BLP IVs) would be available as instruments.

S.4 Semiparametric Models

The previous example considered Z_{it} with large support. One can instead move in the opposite direction to consider Z_{it} with more limited dimension

[7] See our early working paper, Berry and Haile (2010).

and support than required in the text. We do so here by considering semiparametric specifications of inverse demand that generalize parametric models commonly used in practice. We focus on the case of a one-dimensional binary Z_{it}, taking values 0 and 1.

Consider a semiparametric nested logit model where the inverted demand system in each market takes the form

$$g_j(z_{it})+\xi_{jt} = \ln(s_{jt}(z_{it})/s_{0t}(z_{it}))-\theta\ln(s_{j/n,t}(z_{it}))+\alpha p_{jt} \quad j=1,\ldots,J. \tag{S.21}$$

Here we have conditioned on X_t (treating it fully flexibly) and suppressed it from the notation. On the RHS, $s_{jt}(z_{it})$ denotes good j's (observable) choice probability in market t conditional on z_{it}, with $s_{j/n,t}(z_{it})$ denoting the within-nest conditional choice probability. The scalar θ denotes the usual "nesting parameter."

The nested logit model embeds normalizations of the indices and demand function analogous to our choices of $A(x)$ and $B(x)$ in section 2.5. However, we must still normalize the location of either Ξ_{jt} or g_j for each j to pose the identification question. Here we set $g_j(0)=0$ for all j, breaking with our prior convention by leaving each $\mathbb{E}[\Xi_{jt}]$ free.

Here (S.21) implies the two equations

$$g_j(1)+\xi_{jt} = \ln(s_{jt}(1)/s_{0t}(1))-\theta\ln(s_{j/n,t}(1))+\alpha p_{jt} \tag{S.22}$$
$$g_j(0)+\xi_{jt} = \ln(s_{jt}(0)/s_{0t}(0))-\theta\ln(s_{j/n,t}(0))+\alpha p_{jt} \tag{S.23}$$

for every product j and market t. Differencing these equations in one market, we obtain

$$g_j(1)=\ln\left(\frac{s_{jt}(1)}{s_{0t}(1)}\right)-\ln\left(\frac{s_{jt}(0)}{s_{0t}(0)}\right)-\theta\left[\ln(s_{j/n,t}(1))-\ln(s_{j/n,t}(0))\right] \tag{S.24}$$

for $j=1,\ldots,J$. This is J equations in $J+1$ unknowns: θ and $g_1(1),\ldots,g_J(1)$.

Move now to a different market t', where the observed choice probabilities are different (perhaps because $\xi_t \neq \xi_{t'}$). For this market, (S.24) takes the form

$$g_j(1)=\ln\left(\frac{s_{jt'}(1)}{s_{0t'}(1)}\right)-\ln\left(\frac{s_{jt'}(0)}{s_{0t'}(0)}\right)-\theta\left[\ln(s_{j/n,t'}(1))-\ln(s_{j/n,t'}(0))\right] \tag{S.25}$$

for $j=1,\ldots,J$. This provides J new equations with no new unknowns. Given minimal variation in choice probabilities across markets, ensuring that

$$\ln(s_{j/n,t}(1))-\ln(s_{j/n,t}(0)) \neq \ln(s_{j/n,t'}(1))-\ln(s_{j/n,t'}(0)) \tag{S.26}$$

for at least one good j, one can then solve for θ and $g_1(1),\ldots,g_J(1)$. Identification of the remaining parameter α can then be obtained from the "regression" equation (obtained from (S.21))

$$\ln(s_{jt}(z_{it})/s_{0t}(z_{it})) - g_j(z_{it}) - \theta \ln(s_{j/n,t}(z_{it}) = -\alpha p_{jt} - \xi_{jt} \tag{S.27}$$

using a single excluded instrument for price—e.g., an excluded exogenous market-level cost shifter or markup shifter that affects all prices. This compares to the usual requirement of two instruments in the fully parametric nested logit when one has only market-level data (see Berry (1994)). Thus, as in the fully nonparametric case, micro data cuts the number of required instruments by half. The intercept in this regression equation picks up the mean of the unobservable Ξ_{jt}.

Observe that here the argument proceeds in two steps, mirroring those in the main text. We first use a combination of within- and cross-market variation to uncover the function g and consumer substitution patterns (determined here by the parameter θ). We then use cross-market instrumental variables restrictions to separate the roles of prices (and other market-level factors) from the effects of the demand shocks.

Of course, in the first step (S.26) may typically hold for all j, implying overidentification and suggesting the potential to introduce a more flexible specification of the inverse demand mapping—e.g., adding nests or BLP-style random coefficients. Furthermore, there is no reason to limit attention to just two markets in the first step: each additional market adds new equations but no new unknowns.[8] Although this discussion is informal, it suggests the potential to obtain identification of semiparametric demand models with flexible substitution patterns, using (along with instrument(s) for prices) consumer-level characteristics with substantially more limited dimension and support than we required for the fully nonparametric model in the text.

S.5 Beyond Discrete Choice

Although the text emphasizes the case in which the consumer-level quantities Q_{ijt} take the particular form implied by a discrete choice model, nothing in our proofs requires this. In other settings, the demand function $\jmath$ defined in (1) may simply be reinterpreted as the expected vector of quantities demanded conditional on $(Z_{it}, P_t, X_t, \Xi_t)$.[9] Applying our results to continuous demand is

[8]It is easy to see how the example here generalizes if we allow Z_{it} to have more than two points of support. With K_Z points of support, differencing the analogs of (S.22) and (S.23) for one market yields $(K_Z - 1) \times J$ equations in $1 + (K_Z - 1) \times J$ unknowns. Each new market adds $(K_Z - 1) \times J$ equations and no new unknowns in the first step of the argument.

[9]Note that the demand faced by firms in market t is the expectation (over the distribution of consumer-level observables in the market) of this expected demand.

therefore just a matter of verifying the suitability of our assumptions.[10]

As one possibility, consider a "mixed CES" model of continuous choice, similar to the model in Adao, Costinot, and Donaldson (2017), with $J+1$ products. Here we introduce the notation Y_{it} for the observed income of consumer i in market t, measured in units of the numeraire good 0. For this example we treat Y_{it} as an additional consumer-level observable, beyond the J-dimensional Z_{it} assumed to obey our index restrictions. We again focus on the case in which X_t is exogenous.

Each consumer i in market t has utility over consumption vectors $q \in \mathbb{R}_+^{J+1}$ given by

$$u\left(q; z_{it}, p_t, x_t, \xi_t\right) = \left(\sum_{j=0}^{J} \phi_{ijt} q_j^{\rho}\right)^{1/\rho},$$

where $\rho \in (0,1)$ is a parameter and each ϕ_{ijt} represents idiosyncratic preferences of consumer i. Normalizing $\phi_{i0t} = 1$, let

$$\phi_{ijt} = \exp\left[(1-\rho)\left(g_j\left(z_{it}, x_t\right) + \xi_{jt} + x_{jt}\beta_{it}\right)\right],\ j = 1, \ldots, J,$$

where β_{it} is a random vector (with distribution F_β) representing consumer-level preferences for product characteristics. With $p_{0t} = 1$, familiar CES algebra shows that Marshallian demands are

$$q_{ijt} = \frac{y_{it} \exp\left(g_j\left(z_{it}, x_t\right) + \xi_{jt} + x_{jt}\beta_{it} - \alpha \ln(p_{jt})\right)}{1 + \left[\sum_{k=1}^{J} \exp\left(g_k\left(z_{it}, x_t\right) + \xi_{kt} + x_{kt}\beta_{it} - \alpha\rho \ln(p_{kt})\right)\right]}, \tag{S.28}$$

where $\alpha = 1/(1-\rho)$. It is easy to show that our Assumptions 1–3 are satisfied for the expected demand functions

$$\sigma_t(g(z_{it}, x_t) + \xi_t, y_{it}, x_t, p_t) = \mathbb{E}\left[Q_{it} | z_{it}, y_{it}, p_t, x_t, \xi_t\right],$$

where the jth component of $\mathbb{E}[Q_{it}|z_{it}, y_{it}, x_t, p_t, \xi_t]$ is

$$\int \frac{y_{it} \exp\left(g_j\left(z_{it}, x_t\right) + \xi_{jt} + x_{jt}\beta_{it} - \alpha \ln(p_{jt})\right)}{1 + \left[\sum_{k=1}^{J} \exp\left(g_k\left(z_{it}, x_t\right) + \xi_{kt} + x_{kt}\beta_{it} - \alpha\rho \ln(p_{kt})\right)\right]} \, dF_\beta(\beta_{it}).$$

[10] Berry, Gandhi, and Haile (2013) describe a broad class of continuous choice models that can satisfy the key injectivity property of Assumption 2. These can include mixed continuous/discrete settings, where individual consumers may purchase zero or any positive quantity of each good.