Econometrica Supplementary Material

SUPPLEMENT TO "LARGE MARKET ASYMPTOTICS FOR DIFFERENTIATED PRODUCT DEMAND ESTIMATORS WITH ECONOMIC MODELS OF SUPPLY"

(*Econometrica*, Vol. 84, No. 5, September 2016, 1961–1980)

BY TIMOTHY B. ARMSTRONG

This supplement contains proofs as well as auxiliary and Monte Carlo results. Section A contains proofs of results in the main text. Section B discusses large market asymptotics in some additional settings. Section C gives the details of the Monte Carlo study, and presents additional Monte Carlo results for designs not reported in the main text.

A. PROOFS

THIS SECTION PRESENTS PROOFS of the results in the main text. Section A.1 states and proves equivalence results used in the rest of the section, including the proof of Theorem 1 from the main text. The rest of the section contains proofs of the remaining results in the main text.

A.1. *Equivalence Results for IV Estimators*

Many of the results in the paper are based on the IV equivalence results. The results follow from characterizations of the asymptotic behavior of IV estimators under possible lack of identification (this step follows known results in the literature; see, for example, Staiger and Stock (1997)) along with bounds on the difference between sample moments involving different covariates. The following theorems are stated for a general linear IV estimator $\hat{\beta} = [(\frac{1}{J}\sum_{j=1}^J z_j x_j')W_J(\frac{1}{J}\sum_{j=1}^J z_j x_j')]^{-1}(\frac{1}{J}\sum_{j=1}^J z_j x_j')' W_J(\frac{1}{J}\sum_{j=1}^J z_j y_j')$, where z_j is a vector of instruments, x_j is a vector of covariates, and $y_j = x_j'\beta + \xi_j$ (in the notation of the rest of the paper, this theorem is used with (x_j, p_j) taking the place of x_j and $(\alpha, \beta')'$ taking the place of β). In what follows, the behavior of $\hat{\beta}$ under a sequence x_j^* and y_j^* with $y_j^* = x_j^{*\prime}\beta + \xi_j$ is compared to the behavior of β under the original sequences.

ASSUMPTION 1: (i) *For some sequence of $k \times d$ matrices $\{M_{zx,J}\}_{J=1}^\infty$, an invertible $d \times d$ matrix H, and nonnegative integers d_1 and d_2 with $d_1 + d_2 = d$, the first d_1 columns of $M_{zx,J}H$ are* 0 *for all J, and $\sqrt{J}(\frac{1}{J}\sum_{j=1}^J z_j x_j' - M_{zx,J}) \xrightarrow{d} Z_{zx}$ and $M_{zx,J} \to M_{zx}$ for some matrix M_{zx} and a $k \times d$ random matrix Z_{zx} such that the last d_2 columns of $M_{zx}H$ have rank d_2*. (ii) *For a limiting multivariate normal random vector $Z_{z\xi}$ with nonsingular variance, $\frac{1}{\sqrt{J}}\sum_{j=1}^J z_j \xi_j \xrightarrow{d} Z_{z\xi}$ jointly with the convergence in distribution in part* (i). (iii) *We have $W_J \xrightarrow{p} W$ for some positive definite weighting matrix W.*

 DOI: 10.3982/ECTA10600

ASSUMPTION 2: *We have* $\sqrt{J}\max_j \|x_j^* - x_j\| \xrightarrow{p} 0$ *and* $\frac{1}{J}\sum_{j=1}^J \|z_j\| = \mathcal{O}_P(1)$.

THEOREM 5: *Under Assumption 1, the following holds.*

(i) *Define* $T_J = H^{-1}(\hat{\beta} - \beta)$ *with* T_{1J} *the first* d_1 *elements and* T_{2J} *the last* d_2 *elements. Then*

$$\begin{pmatrix} T_{1J} \\ \sqrt{J}T_{2J} \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \left((Z_{zx}H_1)'Q_{W,2}'WQ_{W,2}Z_{zx}H_1\right)^{-1}(Z_{zx}H_1)'Q_{W,2}'WQ_{W,2}Z_{z\xi} \\ \left((M_{zx}H_2)'Q_{W,1}'WQ_{W,1}M_{zx}H_2\right)^{-1}(M_{zx}H_1)'Q_{W,1}'WQ_{W,1}Z_{z\xi} \end{pmatrix},$$

where $H = (H_1, H_2)$ *with* H_1 *forming the first* d_1 *columns and* H_2 *forming the remaining columns,* $Q_{W,1}$ *is the* W *inner product projection matrix for the orthogonal complement of the column span of* $Z_{zx}H_1$*, and* $Q_{W,2}$ *is the* W *inner product projection matrix for the orthogonal complement of the column span of* $M_{zx}H_2$.

(ii) *If Assumption 2 holds as well, then letting* $\hat{\beta}^*$ *be the estimator with* x_j^* *and* y_j^* *replacing* x_j *and* y_j, $\|\hat{\beta} - \hat{\beta}^*\| \xrightarrow{p} 0$.

PROOF: Part (i) essentially follows from applying results for partially identified IV (see, for example, Stock and Wright (2000)) to a version of the model that is reparameterized so that the parameter of interest is $H^{-1}\beta$. We have, letting A_J be the $d \times d$ diagonal matrix with the first d_1 diagonal entries equal to 1 and the last d_2 equal to $\sqrt{J}$,

$$\begin{aligned}\hat{\beta} - \beta &= \left(\left[\sum_{j=1}^J z_j'x_j\right]' W_J \left[\sum_{j=1}^J z_j'x_j\right]\right)^{-1} \\ &\quad \times \left[\sum_{j=1}^J z_j'x_j\right]' W_J \left[\sum_{j=1}^J z_j\left(y_j - x_j'\beta\right)\right] \\ &= \arg\min_\gamma \left\|E_J z\xi - E_J zx'\gamma\right\|_{W_J},\end{aligned}$$

where $E_J z\xi = \frac{1}{J}\sum_{j=1}^J z_j\xi_j$ and $E_J zx' = \frac{1}{J}\sum_{j=1}^J z_j x_j'$. Thus,

$$\begin{aligned}\begin{pmatrix} T_{1J} \\ \sqrt{J}T_{2J} \end{pmatrix} &= A_J H^{-1}(\hat{\beta} - \beta) = \arg\min_\gamma \left\|E_J z\xi - E_J zx'HA_J^{-1}\gamma\right\|_{W_J} \\ &= \arg\min_\gamma \left\|\sqrt{J}E_J z\xi - \sqrt{J}E_J zx'HA_J^{-1}\gamma\right\|_{W_J} \\ &= \arg\min_\gamma \left\|\sqrt{J}E_J z\xi - \left(\sqrt{J}E_J zx'H_1, E_J zx'H_2\right)\gamma\right\|_{W_J}.\end{aligned}$$

By the continuous mapping theorem, this converges to

$$\arg\min_{\gamma}\left\| Z_{z\xi} - (Z_{zx}H_1, M_{zx}H_2)\gamma\right\|_{W_J}.$$

The result follows from applying the partitioned least squares formula to this expression.

For part (ii), note that, under Assumptions 1 and 2, Assumption 1 will also hold with x_j^* and y_j^*. In fact, we will have $(\sqrt{J}(\frac{1}{J}\sum_{j=1}^{J} z_j x_j^{*\prime} - M_{zx,J}), \sqrt{J}(\frac{1}{J}\sum_{j=1}^{J} z_j x_j' - M_{zx,J})) \xrightarrow{d} (Z_{zx}, Z_{zx})$. The result follows by applying the above results to $\hat{\beta}^*$, where we modify the above argument by applying the continuous mapping theorem to $(T_{1J}', \sqrt{J}T_{2J}')' - (T_{1J}^{*\prime}, \sqrt{J}T_{2J}^{*\prime})'$ to show that this quantity converges in distribution (and in probability) to a limiting distribution that can be seen to be identically zero. *Q.E.D.*

Theorem 1 follows by verifying the conditions of Theorem 5.

PROOF OF THEOREM 1: The result follows from Theorem 5 with (x_j, p_j) in place of x_j^*. The first part of Assumption 2 follows from condition (i) in Theorem 1, with the boundedness of $\frac{1}{J}\sum_{j=1}^{J}\|z_j\|$ following from condition (ii), since x_j contains a constant. Assumption 1 follows from condition (ii), with

$$\begin{aligned}
M_{zx,J} &= \frac{1}{J}\sum_{j=1}^{J} E z_j\left(x_j, MC_j + b^*\right) \\
&= \frac{1}{J}\sum_{j=1}^{J} E\begin{pmatrix} x_j \\ h_j(x_{-j})\end{pmatrix}\begin{pmatrix} x_j' & MC_j + b^*\end{pmatrix} \\
&= \frac{1}{J}\sum_{j=1}^{J}\begin{pmatrix} E x_j x_j' & E x_j MC_j + b^* E x_j \\ E h_j(x_{-j}) E x_j' & \left[E h_j(x_{-j})\right]\left(EMC_j + b^*\right)\end{pmatrix} \\
&= \frac{1}{J}\sum_{j=1}^{J}\begin{pmatrix} 1 & 0 & 0 \\ 0 & I_{d_x-1} & 0 \\ E h_j(x_{-j}) & 0 & 0\end{pmatrix} \\
&\quad\times\begin{pmatrix} 1 & E w_j' & EMC_j + b^* \\ E w_j & E w_j w_j' & E w_j MC_j + b^* E w_j \\ 0 & 0 & 1\end{pmatrix},
\end{aligned}$$

where $x_j = (1, w_j')'$, and with $d_1 = 1$ and

$$H = \begin{pmatrix} 1 & E w_j' & EMC_j + b^* \\ E w_j & E w_j w_j' & E w_j MC_j + b^* E w_j \\ 0 & 0 & 1\end{pmatrix}^{-1}\begin{pmatrix} 0 & 0 & 1 \\ 0 & I_{d_X-1} & 0 \\ 1 & 0 & 0\end{pmatrix}$$

(which does not depend on j by the i.i.d. assumption). *Q.E.D.*

The next theorem deals with the case where M_{zx} is full rank, leading to consistent estimators. The theorem uses a slightly weaker version of the assumptions used for Theorem 5 (with M_{zx} full rank).

ASSUMPTION 3: *Assumption 1 holds with part* (i) *replaced by the condition that* $\frac{1}{J}\sum_{j=1}^{J} z_j x_j' \xrightarrow{p} M_{zx}$, *with* M_{zx} *full rank.*

ASSUMPTION 4: *We have* $\max_j \|x_j^* - x_j\| \xrightarrow{p} 0$ *and* $\frac{1}{J}\sum_{j=1}^{J}\|z_j\| = \mathcal{O}_P(1)$.

THEOREM 6: *Under Assumptions 3 and 4,* $\sqrt{J}(\hat\beta - \beta) \xrightarrow{d} (M_{zx}' W M_{zx})^{-1} \times M_{zx}' W Z_{z\xi}$, *and the same holds for* $\sqrt{J}(\hat\beta^* - \beta)$.

PROOF: Under these assumptions, Assumption 3 holds for both the original and the starred quantities (note that $\frac{1}{J}\sum_{j=1}^{J} z_j' x_j^* = \frac{1}{J}\sum_{j=1}^{J} z_j' x_j + \frac{1}{J}\sum_{j=1}^{J} z_j'(x_j^* - x_j)$; the first term converges in probability to M_{zx} and the second term is bounded by $\frac{1}{J}\sum_{j=1}^{J}\|z_j\|\cdot\|x_j^* - x_j\|$, which converges in probability to zero under these assumptions). The result then follows since

$$
\begin{aligned}
\sqrt{J}(\hat\beta - \beta) &= \left(\left[\frac{1}{J}\sum_{j=1}^{J} z_j' x_j\right]' W_J \left[\frac{1}{J}\sum_{j=1}^{J} z_j' x_j\right]\right)^{-1} \\
&\quad \times \left[\frac{1}{J}\sum_{j=1}^{J} z_j' x_j\right]' W_J \left[\frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j\big(y_j - x_j'\beta\big)\right] \\
&\xrightarrow{d} \big(M_{zx}' W M_{zx}\big)^{-1} M_{zx}' W Z_{z\xi}
\end{aligned}
$$

by the continuous mapping theorem since $\frac{1}{J}\sum_{j=1}^{J} z_j' x_j \xrightarrow{p} M_{zx}$ and $\frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j(y_j - x_j'\beta) = \frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j\xi_j \xrightarrow{d} Z_{z\xi}$, and, for the starred quantities,

$$
\begin{aligned}
\sqrt{J}\big(\hat\beta^* - \beta\big) &= \left(\left[\frac{1}{J}\sum_{j=1}^{J} z_j' x_j^*\right]' W_J \left[\frac{1}{J}\sum_{j=1}^{J} z_j' x_j^*\right]\right)^{-1} \\
&\quad \times \left[\frac{1}{J}\sum_{j=1}^{J} z_j' x_j^*\right]' W_J \left[\frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j\big(y_j^* - x_j^{*\prime}\beta\big)\right] \\
&\xrightarrow{d} \big(M_{zx}' W M_{zx}\big)^{-1} M_{zx}' W Z_{z\xi}
\end{aligned}
$$

since $\frac{1}{J}\sum_{j=1}^{J} z_j' x_j^* \xrightarrow{p} M_{zx}$ and $\frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j(y_j^* - x_j^{*\prime}\beta) = \frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j\xi_j \xrightarrow{d} Z_{z\xi}$.

Q.E.D.

Note that the conditions for Theorem 5 require $\sqrt{J}\max_j \|x_j^* - x_j\| \xrightarrow{p} 0$, while Theorem 6 requires only the weaker condition that $\max_j \|x_j^* - x_j\| \xrightarrow{p} 0$. This is because the asymptotic distribution of $\hat{\beta}$ in Theorem 5 depends on the asymptotic distribution of $\sqrt{J}(\frac{1}{J}\sum_{j=1}^J z_j x_j' - M_{zx,J})$, while the asymptotic distribution of $\sqrt{J}(\hat{\beta} - \beta)$ in Theorem 6 only uses the fact that $\frac{1}{J}\sum_{j=1}^J z_j x_j' \xrightarrow{p} M_{zx}$, and does not depend on its asymptotic distribution. To get the same results with the starred quantities, weaker conditions suffice in the case of Theorem 6.

A.2. *Proof of Theorem 2*

We prove a slightly more general version of Theorem 2 with the boundedness condition on MC_j, x_j, and ξ_j generalized to an exponential tail condition. In particular, we replace the condition that (x_j, ξ_j, MC_j) is bounded with the following condition: for some constants C and $\varepsilon > 0$, $P(|\xi_j| \geq t) \leq C\exp(-\varepsilon t^{1+\varepsilon})$, $P(|MC_j| \geq t) \leq C\exp(-\varepsilon t^{1+\varepsilon})$, and $P(\|x_j\| \geq t) \leq C\exp(-\varepsilon t^{2+\varepsilon})$.

Before proceeding to the proof, note that, formally, the theorem applies to the triangular array of prices $p_{j,J}$ arising from any sequence of Nash–Bertrand equilibria (defined for any realization of the primitives $\{x_j, MC_j, \xi_j\}_{j=1}^J$), defined for each J. This can be made explicit by writing prices as a function of $\{x_j, MC_j, \xi_j\}_{j=1}^J$, and an additional random variable ω_J that determines equilibrium selection in the case of multiple equilibria (which may be arbitrarily correlated with the remaining variables): $p_j = p_{j,J}(\{x_j, MC_j, \xi_j\}_{j=1}^J, \omega_J)$. We assume that an equilibrium exists on a probability 1 set of $\{x_j, MC_j, \xi_j\}_{j=1}^J$ for each J.

The proof of the theorem uses only the first order condition for each firm's best response, which holds regardless of how ω_J determines the equilibrium in the case of multiple equilibria (since the strategy space for prices is $(-\infty, \infty)$, the best response problem must be maximized at an interior solution in equilibrium). Thus, to simplify notation, we leave the dependence of prices on ω_J implicit and write p_j instead of $p_{j,J}(\{x_j, MC_j, \xi_j\}_{j=1}^J, \omega_J)$ in the remainder of the proof. To further simplify notation in the proof of this theorem, we also define $g_j(\zeta) = x_j'(\beta + \zeta) + \xi_j$.

At several places in the proof, we use the fact that, for some constant K, $\int \exp(a\|\zeta\|)\, dP(\zeta) \leq K\exp(Ka^2)$ for any a and $\int \exp(t'\zeta)\, dP(\zeta) \leq \int \exp(\|t\| \cdot \|\zeta\|)\, dP(\zeta) \leq K\exp(K\|t\|^2)$ for any t. This follows since the tails of ζ are bounded by the tails of the normal random variable (for $Z \sim N(\mu, \Sigma)$, $E\exp(a\|Z\|)$ is bounded by $K\exp(Ka^2)$ for a constant K that depends on μ and σ^2).

Rearranging the markup formula for firm j gives

$$p_j - MC_j - \frac{1}{\alpha} = \frac{\displaystyle\int \tilde{s}_j^2(\delta, \zeta)\, dP_\zeta(\zeta)}{\alpha \displaystyle\int \tilde{s}_j(\delta, \zeta)\big(1 - \tilde{s}_j(\delta, \zeta)\big)\, dP_\zeta(\zeta)}$$

$$= \frac{1}{\alpha} \frac{\dfrac{\displaystyle\int \tilde{s}_j^2(\delta, \zeta)\, dP_\zeta(\zeta)}{\displaystyle\int \tilde{s}_j(\delta, \zeta)\, dP_\zeta(\zeta)}}{1 - \dfrac{\displaystyle\int \tilde{s}_j^2(\delta, \zeta)\, dP_\zeta(\zeta)}{\displaystyle\int \tilde{s}_j(\delta, \zeta)\, dP_\zeta(\zeta)}}.$$

Thus, it suffices to show that $\frac{\int \tilde{s}_j^2(\delta,\zeta)\, dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\, dP_\zeta(\zeta)}$ converges to zero more quickly than $1/\sqrt{J}$ uniformly over $1 \le j \le J$. To this end, we prove the following lemma.

LEMMA 1: *Suppose that, for some constants B_J and A_J, $|\xi_j| \le A_J$, $|p_j| \le A_J$ and $\|x_j\| \le B_J$ for all j, and that the tails of ζ are bounded by the tails of a normal variable. Then, for some constant C,*

$$\frac{\displaystyle\int \tilde{s}_j^2(\delta, \zeta)\, dP_\zeta(\zeta)}{\displaystyle\int \tilde{s}_j(\delta, \zeta)\, dP_\zeta(\zeta)} \le \frac{1}{J} C \exp\bigl(C \cdot A_J + C \cdot B_J^2\bigr).$$

PROOF: Under these conditions,

$$\begin{aligned}
\tilde{s}_j(\delta, \zeta) &\le \frac{\exp\bigl(B_J \|\beta + \zeta\| + \alpha A_J + A_J\bigr)}{\displaystyle\sum_{k=1}^{J} \exp\bigl(-B_J \|\beta + \zeta\| - A_J\bigr)} \\
&= \frac{1}{J} \exp\bigl(2B_J \|\beta + \zeta\| + (\alpha + 2) A_J\bigr).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\tilde{s}_j(\delta, \zeta) &\ge \frac{\exp\bigl(-B_J \|\beta + \zeta\| - A_J\bigr)}{\displaystyle\sum_{k=1}^{J} \exp\bigl(B_J \|\beta + \zeta\| + \alpha A_J + A_J\bigr)} \\
&= \frac{1}{J} \exp\bigl(-2B_J \|\beta + \zeta\| - (\alpha + 2) A_J\bigr).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\frac{\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)} \\
&\quad\le \frac{\int \frac{1}{J^2}\exp\bigl(2\cdot 2B_J\|\beta+\zeta\|+2(\alpha+2)A_J\bigr)\,dP_\zeta(\zeta)}{\int \frac{1}{J}\exp\bigl(-2B_J\|\beta+\zeta\|-(\alpha+2)\cdot A_J\bigr)\,dP_\zeta(\zeta)} \\
&\quad= \frac{1}{J}\exp\bigl(3(\alpha+2)A_J\bigr)\frac{\int \exp\bigl(4B_J\|\beta+\zeta\|\bigr)\,dP_\zeta(\zeta)}{\int \exp\bigl(-2B_J\|\beta+\zeta\|\bigr)\,dP_\zeta(\zeta)}.
\end{aligned}$$

The result follows since the integral in the denominator is bounded from below by $\exp(-2\cdot B_J\cdot K)\cdot P_\zeta(\|\zeta\|\le K)\ge \exp(-2\cdot B_J\cdot K)\cdot(1/2)$ for large enough K, and the numerator is bounded from above by a constant times $K\exp(KB_J^2)$ for large enough K. *Q.E.D.*

The bound in Lemma 1 will decrease more quickly than $1/\sqrt{J}$ as long as $A_J/\log J\to 0$ and $B_J^2/\log J\to 0$. For the bounds on the primitives x_j, ξ_j, and MC_j, this follows easily from the bounds on tail probabilities, as shown in the next lemma. For the bound on prices, a more involved argument is needed, which constitutes the remainder of the proof.

LEMMA 2: *Let $\{u_j\}_{j=1}^J$ be a sequence of random variables such that $P(u_j\ge t)\le C\exp(-t^\gamma/C)$ for some γ and C that do not depend on j. Then, for any $\varepsilon>0$,*

$$P\Bigl(\max_{1\le j\le J} u_j\ge (C+\varepsilon)^{1/\gamma}(\log J)^{1/\gamma}\Bigr)\to 0.$$

PROOF: We have

$$\begin{aligned}
P\Bigl(\max_{1\le j\le J} u_j\ge (C+\varepsilon)^{1/\gamma}(\log J)^{1/\gamma}\Bigr) &\le \sum_{j=1}^J P\bigl(u_j\ge (C+\varepsilon)^{1/\gamma}(\log J)^{1/\gamma}\bigr) \\
&\le J\cdot C\exp\bigl(-(C+\varepsilon)(\log J)/C\bigr) \\
&= J\cdot C\cdot J^{-(C+\varepsilon)/C}\overset{J\to\infty}{\to} 0. \qquad \textit{Q.E.D.}
\end{aligned}$$

It follows from Lemma 2 that all of the conditions of Lemma 1, except for the bound on price, hold with probability 1 for $A_J=C(\log J)^{1-\varepsilon}$ and $B_J=$

$C(\log J)^{1/2-\varepsilon}$ for C large enough and ε small enough. To prove the theorem, it suffices to show that $\max_{1\le j\le J}|p_j|$ is also bounded by A_J for C large enough and ε small enough. This follows from the next two lemmas.

LEMMA 3: *Suppose that, for some K_J,*

$$\frac{\int \exp\big(g_j(\zeta)\big)\tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)} \le K_J.$$

Then

$$p_j \le \max\left\{MC_j + \frac{2}{\alpha}, \frac{1}{\alpha}[\log 2 + \log K_J]\right\}.$$

PROOF: Note that $\tilde{s}_j(\delta,\zeta) = \frac{\exp(g_j(\zeta)-\alpha p_j)}{\sum_{k=0}^J \exp(g_k(\zeta)-\alpha p_k)} \le \exp(g_j(\zeta)-\alpha p_j)$, since one of the terms in the denominator is the outside good, with utility 0. Thus,

$$\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta) \le \exp(-\alpha p_j)\int \exp\big(g_j(\zeta)\big)\tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta).$$

Suppose that $\exp(-\alpha p_j)K_J \le 1/2$. Then

$$\begin{aligned} p_j - MC_j - \frac{1}{\alpha} &= \frac{1}{\alpha}\frac{\dfrac{\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}}{1-\dfrac{\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}} \\ &\le \frac{1}{\alpha}\frac{\exp(-\alpha p_j)K_J}{1-\exp(-\alpha p_j)K_J} \le \frac{1}{\alpha}, \end{aligned}$$

where the inequalities follow since $t/(1-t)$ is increasing in t for $0 \le t < 1$. Thus, either $p_j - MC_j - \frac{1}{\alpha} \le \frac{1}{\alpha}$ or

$$\begin{aligned} \exp(-\alpha p_j)K_J > 1/2 \quad &\Longrightarrow \quad -\alpha p_j + \log K_J > \log(1/2) \\ &\Longrightarrow \quad -\alpha p_j > \log(1/2) - \log K_J \\ &\Longrightarrow \quad p_j < -\big[\log(1/2)\big]/\alpha + [\log K_J]/\alpha, \end{aligned}$$

giving the desired bound. *Q.E.D.*

LEMMA 4: *For some constant C,*

$$\frac{\int \exp\big(g_j(\zeta)\big)\tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)} \le C\exp\Big(C\max_{1\le k\le J}\|x_k\|^2 + C\max_{1\le k\le J}\|\xi_k\|\Big).$$

PROOF: We have

$$\begin{aligned}
&\frac{\int \exp\big(g_j(\zeta)\big)\tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)} \\
&\quad= \frac{\displaystyle\int \exp\big(g_j(\zeta)\big)\frac{\exp\big(g_j(\zeta)-\alpha p_j\big)}{\sum_k \exp\big(g_k(\zeta)-\alpha p_k\big)}\,dP_\zeta(\zeta)}{\displaystyle\int \frac{\exp\big(g_j(\zeta)-\alpha p_j\big)}{\sum_k \exp\big(g_k(\zeta)-\alpha p_k\big)}\,dP_\zeta(\zeta)} \\
&\quad= \frac{\displaystyle\int \frac{\exp\big(2g_j(\zeta)\big)}{\sum_k a_k\exp\big(g_k(\zeta)\big)}\,dP_\zeta(\zeta)}{\displaystyle\int \frac{\exp\big(g_j(\zeta)\big)}{\sum_k a_k\exp\big(g_k(\zeta)\big)}\,dP_\zeta(\zeta)},
\end{aligned}$$

where $a_k = \exp(-\alpha p_k)/\sum_\ell \exp(-\alpha p_\ell)$ so that $\sum_{k=1}^J a_k = 1$. By Jensen's inequality, this is bounded by

$$\begin{aligned}
&\left[\int \frac{\exp\big(2g_j(\zeta)\big)}{\sum_k a_k\exp\big(g_k(\zeta)\big)}\,dP_\zeta(\zeta)\right]\left[\int \frac{\sum_k a_k\exp\big(g_k(\zeta)\big)}{\exp\big(g_j(\zeta)\big)}\,dP_\zeta(\zeta)\right] \qquad (6)\\
&\quad\le \left[\int \exp\big(2g_j(\zeta)\big)\sum_k a_k\exp\big(-g_k(\zeta)\big)\,dP_\zeta(\zeta)\right] \\
&\qquad\times\left[\int \frac{\sum_k a_k\exp\big(g_k(\zeta)\big)}{\exp\big(g_j(\zeta)\big)}\,dP_\zeta(\zeta)\right],
\end{aligned}$$

where the last inequality follows from Jensen's inequality applied to $\sum_k a_k \exp(g_k(\zeta))$. We have, for each j and k,

$$\begin{aligned}
&\int \exp\bigl(2g_j(\zeta)\bigr)\exp\bigl(-g_k(\zeta)\bigr)\,dP_\zeta(\zeta) \\
&\quad= \int \exp\bigl(2g_j(\zeta)-g_k(\zeta)\bigr)\,dP_\zeta(\zeta) \\
&\quad= \int \exp\bigl(2x_j'(\beta+\zeta)+2\xi_j-x_k'(\beta+\zeta)-\xi_k\bigr)\,dP_\zeta(\zeta) \\
&\quad= \exp\bigl((2x_j-x_k)'\beta+2\xi_j-\xi_k\bigr) \\
&\qquad\times \int \exp\bigl((2x_j-x_k)'\zeta\bigr)\,dP_\zeta(\zeta).
\end{aligned}$$

Since the tails of ζ are bounded by the tails of a normal variable, this is bounded, for some constant C, by

$$\exp\bigl((2x_j-x_k)'\beta+2\xi_j-\xi_k\bigr)\cdot C\exp\bigl(C\|2x_j-x_k\|^2\bigr).$$

By making C larger, this can be bounded by $C\exp(C\max\{\|x_j\|^2,\|x_k\|^2\}+C\max\{\|\xi_j\|,\|\xi_k\|\})$. Similarly, $\int\exp(g_k(\zeta))/\exp(g_j(\zeta))\,dP_\zeta(\zeta)$ can also be bounded by $C\exp(C\max\{\|x_j\|^2,\|x_k\|^2\}+C\max\{\|\xi_j\|,\|\xi_k\|\})$ for large enough C. Thus, (6) can be bounded by

$$\begin{aligned}
&C^2\Biggl[\sum_{j=1}^{J} a_k C\exp\bigl(C\max\bigl\{\|x_j\|^2,\|x_k\|^2\bigr\}+C\max\bigl\{\|\xi_j\|,\|\xi_k\|\bigr\}\bigr)\Biggr]^2 \\
&\quad\le C^2\exp\Bigl(2C\max_{1\le k\le J}\|x_k\|^2+2C\max_{1\le k\le J}\|\xi_k\|\Bigr).
\end{aligned}$$

The result follows by redefining C. *Q.E.D.*

Putting these lemmas together, it follows that, for some constant C_1,

$$\max_{1\le j\le J}\|p_j\|\le C_1+C_1\max_{1\le j\le J}\|MC_j\|+C_1K\max_{1\le j\le J}\|\xi_j\|+C_1\max_{1\le j\le J}\|x_j\|^2.$$

From this and Lemma 2, it follows that, for C_2 large enough and $\varepsilon>0$ small enough, the conditions of Lemma 1 hold with $A_J=C_2(\log J)^{1-\varepsilon}$ and $B_J=(\log J)^{1/2-\varepsilon}$ with probability approaching 1, giving the desired result.

A.3. *Proof of Theorem 3*

The result follows by verifying the conditions of Theorem 6. First, note that

$$\begin{aligned}
&\frac{1}{J}\sum_{j=1}^{J} z_j\big(x_j', p_j^*\big) - M_{zx} \\
&\quad = \sum_{f\in\mathcal{F}} \frac{|\mathcal{F}_f|}{J}\frac{1}{|\mathcal{F}_f|}\sum_{j\in\mathcal{F}_f}\big[z_j\big(x_j', MC_j + b_f^*\big) - Ez_j\big(x_j', MC_j + b_f^*\big)\big] \\
&\qquad + \frac{1}{J}\sum_{j=1}^{J} Ez_j\big(x_j', p_j^*\big) - M_{zx},
\end{aligned}$$

which converges in probability to zero by the assumption that $\frac{1}{J}\sum_{j=1}^{J} Ez_j(x_j', p_j^*) \to M_{zx}$ and the law of large numbers applied to the i.i.d. sum within each firm f. This verifies the first part of Assumption 3. To verify part (ii) of this assumption, note that

$$\begin{aligned}
\frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j\xi_j &= \sum_{f=1}^{F}\frac{\sqrt{|\mathcal{F}_f|}}{\sqrt{J}}\frac{1}{\sqrt{|\mathcal{F}_f|}}\sum_{j\in\mathcal{F}_f} z_j\xi_j \\
&= \sum_{f=1}^{F}\frac{\sqrt{|\mathcal{F}_f|}}{\sqrt{J}}\frac{1}{\sqrt{|\mathcal{F}_f|}}\sum_{j\in\mathcal{F}_f}\begin{pmatrix} x_j \\ \pi_f\mu_{h,f}\end{pmatrix}\xi_j + R_J,
\end{aligned}$$

where R_J is a vector with the first d_x rows equal to zero, and the remaining d_h rows given by $\frac{1}{\sqrt{J}}\sum_{f=1}^{F}[\frac{1}{J}\sum_{k\in\mathcal{F}_f}\tilde{h}(x_k) - \pi_f\mu_{h,f}][\sum_{j\in\mathcal{F}_f}\xi_j]$ (where d_x is the dimension of x_j and d_h is the dimension of $\tilde{h}(x_k)$). The first term in the display converges to a normal variable with mean zero and variance $\sum_{f=1}^{F}\pi_f V_f$ by the central limit theorem, and $R_J \stackrel{p}{\to} 0$ by the law of large numbers applied to $\frac{1}{J}\sum_{k\in\mathcal{F}_f}\tilde{h}(x_k)$ and the central limit theorem applied to $\frac{1}{\sqrt{J}}\sum_{j\in\mathcal{F}_f}\xi_j$.

To verify Assumption 4, it suffices to show $\max_j |p_j - p_j^*| \stackrel{p}{\to} 0$. Arguing as in Konovalov and Sandor (2010), it can be seen that equation (5) has a unique solution, and defines b as a $\mathbb{R}^F$-valued function that is continuous at $(\pi_0\mu_{r,0}, \pi_1\mu_{r,1}, \dots, \pi_F\mu_{r,F})$ (where $\pi_0 = \lim_{n\to\infty}\hat{\pi}_0 = \lim_{n\to\infty} 1/J = 0$). The difference between p_j and p_j^* can then be written as, for f the firm producing product j, $b_f(\pi_0\mu_{r,0}, \pi_1\mu_{r,1}, \dots, \pi_F\mu_{r,F}) - b_f(\hat{\pi}_0\bar{r}_0, \hat{\pi}_1\bar{r}_1, \dots, \hat{\pi}_F\bar{r}_F)$, which converges in probability to zero by the law of large numbers and the continuous mapping theorem. Since $\max_j |p_j - p_j^*| = \max_f |b_f(\pi_0\mu_{r,0}, \pi_1\mu_{r,1}, \dots, \pi_F\mu_{r,F}) -$

$b_f(\hat{\pi}_0\bar{r}_0, \hat{\pi}_1\bar{r}_1, \ldots, \hat{\pi}_F\bar{r}_F)|$ and the number of firms does not increase with J, the result follows.

A.4. *Proof of Theorem 4*

The following notation is used throughout this section. Let $d_z = d_x + d_h$ (where d_x and d_h are the dimensions of $x_{i,j}$ and $h(x_{i,j})$, respectively). Define $m_2 = \frac{1}{N}\sum_{i=1}^{N}(J_i/\bar{J})^2$, $\tilde{m}_2 = \frac{1}{N}\sum_{i=1}^{N}J_i(J_i - 1)/\bar{J}^2 = m_2 - 1/\bar{J}$, $m_3 = \frac{1}{N}\sum_{i=1}^{N}(J_i/\bar{J})^3$, $m_{2,\infty} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}(J_i/\bar{J})^2$, and $m_{3,\infty} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}(J_i/\bar{J})^3$. Let $r_{i,j} = \exp(x_{i,j}'\beta - \alpha MC_{i,j} - 1 + \xi_{i,j})$. Let $w_{i,j}$ be the nonconstant part of $x_{i,j}$ so that $x_{i,j} = (1, w_{i,j}')'$ and let $\mu_h = E\tilde{h}(x_{i,j})$ $\mu_r = E(r_{i,j})$, $\mu_{xr} = E(x_{i,j}r_{i,j})$, and $\mu_{wr} = E(w_{i,j}r_{i,j})$. For an $n \times m$ matrix M with i, jth entry $M_{i,j}$, $\mathrm{vec}(M)$ is the vectorization of M given by the $ij \times 1$ vector $(M_{1,1}, M_{2,1}, \ldots, M_{n,1}, M_{1,2}, M_{2,2}, \ldots, M_{n,2}, \ldots, M_{1,m}, M_{2,m}, \ldots, M_{n,m})'$. For a $1 \times k$ row vector a, $\mathrm{diag}(a)$ is the $k \times k$ matrix with all off diagonal elements equal to zero and $j \times j$th entry given by the jth element of a. The notation $0_{a\times b}$ and $1_{a\times b}$ is used to denote an $a \times b$ matrix of 0's and 1's, respectively, with the subscript being dropped in cases where the dimension of the matrix of 0's or 1's is clear from the context.

It will be useful to define some additional quantities to describe the asymptotic distribution. Let

$$W_{i,j} = \begin{pmatrix} x_{i,j}x_{i,j}' - Ex_{i,j}x_{i,j}' & x_{i,j}(MC_{i,j} + 1/\alpha) - Ex_{i,j}(MC_{i,j} + 1/\alpha) \\ \mu_h(x_{i,j} - \mu_x)' + \big(\tilde{h}(x_{i,j}) - \mu_h\big)\mu_x' & \mu_h\big(MC_{i,j} - (EMC_{i,j})\big)' + \big(\tilde{h}(x_{i,j}) - \mu_h\big)(EMC_{i,j} + 1/\alpha) \end{pmatrix}$$

and let $u_{i,j} = \xi_{i,j}(x_{i,j}', \mu_h')'$. Let Σ_{Wu} be the variance matrix of $(\mathrm{vec}(W_{i,j})', u_{i,j}')'$ and let $\widetilde{\Sigma}_{Wu}$ be defined by starting with Σ_{Wu} and multiplying diagonal elements $d_x + 1$ through d_z, $d_z + d_x + 1$ through $2d_z$, $2d_z + d_x + 1$ through $3d_z, \ldots$ (those corresponding to the last d_h rows of $W_{i,j}$ and $u_{i,j}$) by $m_{3,\infty}$, and multiplying off diagonal elements in these rows and columns by $m_{2,\infty}$. Define

$$M_1 = \begin{pmatrix} I_{d_x} & 0_{d_x\times 1} \\ 0_{1\times d_x} & 0_{1\times 1} \end{pmatrix},$$

$$H = \begin{pmatrix} E\big(x_{i,j}x_{i,j}'\big) & E\left(x_{i,j}\left(MC_{i,j} + \dfrac{1}{\alpha}\right)\right) \\ 0_{1\times d_x} & 1_{1\times 1} \end{pmatrix},$$

and

$$K_{\tilde{m}_2} = \begin{pmatrix} 1_{1\times 1} & 0 & 0 \\ 0 & I_{d_x-1} & 0 \\ \tilde{m}_2\mu_h & 0 & 1_{d_h\times 1} \end{pmatrix}$$

and let $K_{m_{2,\infty}}$ be defined in the same way, but with $m_{2,\infty}$ replacing $\tilde{m}_2$. Let

$$M_2 = \begin{pmatrix} & 1/\alpha \\ 0_{d_z\times d_x} & \mu_{wr}/(\mu_r\alpha) \\ & \mu_h/\alpha \end{pmatrix}, \qquad A_q = \begin{pmatrix} I_{d_x\times d_x} & 0 \\ 0 & q \end{pmatrix}$$

for any positive real number q. Let

$$\tilde{M}_1 = \begin{pmatrix} 1 & E(w_{i,j})' & E(MC_{i,j}) + \dfrac{1}{\alpha} \\ E(w_{i,j}) & E\big(w_{i,j}w_{i,j}'\big) & E(w_{i,j}MC_{i,j}) + \dfrac{1}{\alpha}E(w_{i,j}) \\ \mu_h\tilde{m}_2 & \mu_h\tilde{m}_2E(w_{i,j})' & \mu_h\tilde{m}_2\left[E(MC_{i,j}) + \dfrac{1}{\alpha}\right] \end{pmatrix}.$$

Note that $A_q^{-1} = A_{1/q}$ and, with the above notation, $\tilde{M}_1 = K_{\tilde{m}_2}M_1H$.

To describe the asymptotic distribution, let $Z_{z\xi}$ be a random vector on $\mathbb{R}^{d_z}$ and let Z_{zx} be a random $d_z \times (d_x + 1)$ matrix, defined on the same probability space such that $(\mathrm{vec}(Z_{zx})', Z_{z\xi}') \sim N(0, \tilde{\Sigma}_{Wu})$. Let $Q = K_{m_{2,\infty}}M_1 + M_2$ and let $\tilde{Q}_{\infty,c} = K_{m_{2,\infty}}M_1 + Z_{zx}H^{-1}\,\mathrm{diag}(0,\dots,0,1) + M_2A_{\sqrt{c}}$ (note that Q is full rank if and only if (iff) $m_{2,\infty} \neq 1$, so that this matrix is full rank under the conditions of the theorem). In the case where $N/\bar{J} \to \infty$, we will show that

$$\begin{aligned}&\mathrm{diag}(\sqrt{\bar{J}N},\dots,\sqrt{\bar{J}N},\sqrt{N/\bar{J}})H\big[\big(\hat{\beta}', -\hat{\alpha}\big)' - \big(\beta', -\alpha\big)'\big] \\ &\quad\xrightarrow{d} (Q'WQ)^{-1}Q'WZ_{z\xi}.\end{aligned}$$

In the case where $N/\bar{J} \to c$ for some finite constant c, we will show that

$$\begin{aligned}&\mathrm{diag}(\sqrt{\bar{J}N},\dots,\sqrt{\bar{J}N},1)H\big[\big(\hat{\beta}', -\hat{\alpha}\big)' - \big(\beta', -\alpha\big)'\big] \\ &\quad\xrightarrow{d} \big(\tilde{Q}_{\infty,c}'W\tilde{Q}_{\infty,c}\big)^{-1}\tilde{Q}_{\infty,c}'WZ_{z\xi}.\end{aligned}$$

We first prove the following lemma.

LEMMA 5: *Under the conditions of Theorem* 4,

$$R_{i,j} \equiv J_i^2\left(p_{i,j} - MC_{i,j} - \frac{1}{\alpha} - \frac{r_{i,j}}{\alpha\displaystyle\sum_{k=1}^{J_i} r_{i,k}}\right)$$

is bounded uniformly over i *as* N *and the* J_i*'s increase*.

PROOF: First, note that

$$p_{i,j} - MC_{i,j} - \frac{1}{\alpha} = \frac{1}{\alpha}\left(\frac{1}{1-s_{i,j}} - 1\right) = \frac{1}{\alpha}\frac{\exp\left(x_{i,j}'\beta - \alpha p_{i,j} + \xi_{i,j}\right)}{\sum_{k\neq j}\exp\left(x_{i,k}'\beta - \alpha p_{i,k} + \xi_{i,k}\right)}. \tag{7}$$

From this formula and the fact that one of the terms in the denominator is the outside good with mean utility zero, it follows that $0 \leq p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \leq \exp(x_{i,j}'\beta - \alpha p_{i,j} + \xi_{i,j}) \leq \exp(x_{i,j}'\beta + \xi_{i,j})$, so that prices are bounded uniformly over i and j. From this and the boundedness of $x_{i,j}$ and $\xi_{i,j}$, it follows that $0 \leq p_{i,j} - MC_{i,j} - 1/\alpha \leq C/(J_i\alpha)$ for C large enough.

Substituting this bound back into (7), we see that (7) is bounded from above by

$$\frac{1}{\alpha}\frac{\exp\left(x_{i,j}'\beta - \alpha MC_{i,j} - 1 + \xi_{i,j}\right)}{\sum_{k\neq j}\exp\left(x_{i,k}'\beta - \alpha MC_{i,k} - 1 - C/J_i + \xi_{i,k}\right)}$$
$$= \frac{1}{\alpha}\frac{\exp\left(x_{i,j}'\beta - \alpha MC_{i,j} - 1 + \xi_{i,j}\right)}{\exp(-C/J_i)\sum_{k\neq j}\exp\left(x_{i,k}'\beta - \alpha MC_{i,k} - 1 + \xi_{i,k}\right)}$$

and from below by

$$\frac{1}{\alpha}\frac{\exp\left(x_{i,j}'\beta - \alpha MC_{i,j} - 1 - C/J_i + \xi_{i,j}\right)}{\sum_{k\neq j}\exp\left(x_{i,k}'\beta - \alpha MC_{i,j} - 1 + \xi_{i,k}\right)}.$$

Thus,

$$\exp(-C/J_i)\frac{r_{i,j}}{\alpha\sum_{k\neq j} r_{i,k}} \leq p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \leq \exp(C/J_i)\frac{r_{i,j}}{\alpha\sum_{k\neq j} r_{i,k}}.$$

Using the fact that, for a constant C_1 that depends only on C, $\exp(C/J_i) \leq 1 + C_1/J_i$ and $\exp(-C/J_i) \geq 1 - C_1/J_i$, we have

$$(1 - C_1/J_i)\frac{r_{i,j}}{\alpha\sum_{k\neq j} r_{i,k}} \leq p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \leq (1 + C_1/J_i)\frac{r_{i,j}}{\alpha\sum_{k\neq j} r_{i,k}}.$$

Thus,

$$\left| p_{i,j} - MC_{i,j} - \frac{1}{\alpha} - \frac{r_{i,j}}{\alpha \sum_{k \neq j} r_{i,k}} \right| \leq \frac{C_1}{J_i} \frac{r_{i,j}}{\alpha \sum_{k \neq j} r_{i,k}} \leq \frac{C_1}{J_i(J_i - 1)} \frac{\bar{r}}{\alpha \underline{r}},$$

where $\bar{r}$ and $\underline{r}$ are positive upper and lower bounds for $r_{i,j}$ (which exist by boundedness of $x_{i,j}$, $\xi_{i,j}$, and $MC_{i,j}$). The result now follows by using the triangle inequality and noting that

$$\begin{aligned} \left| \frac{r_{i,j}}{\alpha \sum_{k=1}^{J_i} r_{i,k}} - \frac{r_{i,j}}{\alpha \sum_{k \neq j} r_{i,k}} \right| &= \frac{r_{i,j}^2}{\alpha \left(\sum_{k=1}^{J_i} r_{i,k} \right) \left(\sum_{k \neq j} r_{i,k} \right)} \\ &\leq \frac{\bar{r}^2}{\alpha \underline{r}^2 J_i (J_i - 1)}. \end{aligned}$$

Q.E.D.

This result is used in the following lemmas, which concern the sample means involved in the IV estimator.

LEMMA 6: *Under the conditions of Theorem* 4,

$$\frac{1}{N\bar{J}} \sum_{i=1}^{N} \sum_{j=1}^{J_i} \left[\frac{1}{\bar{J}} \sum_{k \neq j} \tilde{h}(x_{i,k}) \right] \left(p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \right) = \frac{1}{\bar{J}} \left(\frac{\mu_h}{\alpha} + o_P(1) \right).$$

PROOF: The term that is claimed to be $o_P(1)$ is given by

$$\begin{aligned} (8) \qquad & \bar{J} \left\{ \frac{1}{N\bar{J}} \sum_{i=1}^{N} \sum_{j=1}^{J_i} \left[\frac{1}{\bar{J}} \sum_{k \neq j} \tilde{h}(x_{i,k}) \right] \left(p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \right) - \frac{1}{\bar{J}} \frac{\mu_h}{\alpha} \right\} \\ &= \frac{1}{N\bar{J}} \left\{ \sum_{i=1}^{N} \sum_{j=1}^{J_i} \left[\sum_{k=1}^{J_i} \tilde{h}(x_{i,k}) \right] \left(p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \right) - \frac{\mu_h}{\alpha} N\bar{J} \right\} \\ &\quad - \frac{1}{N\bar{J}} \sum_{i=1}^{N} \sum_{j=1}^{J_i} \tilde{h}(x_{i,j}) \left(p_{i,j} - MC_{i,j} - \frac{1}{\alpha} \right). \end{aligned}$$

The last term goes to zero, since it is bounded by $\frac{1}{N\bar{J}} \sum_{i=1}^{N} \sum_{j=1}^{J_i} |\tilde{h}(x_{i,j})| \cdot C/J_i$, where C/J_i is a bound for $|p_{i,j} - MC_{i,j} - \frac{1}{\alpha}|$ (where the existence of such a bound follows from Lemma 5). Using Lemma 5, the first term in (8) is equal to

$$\frac{1}{N\bar{J}}\left\{\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left[\sum_{k=1}^{J_i}\tilde{h}(x_{i,k})\right]\left(\frac{r_{i,j}}{\alpha\sum_{k=1}^{J_i}r_{i,k}}+\frac{R_{i,j}}{J_i^2}\right)-\frac{\mu_h}{\alpha}N\bar{J}\right\}$$
$$=\frac{1}{N\bar{J}}\left\{\sum_{i=1}^{N}\left[\sum_{j=1}^{J_i}\tilde{h}(x_{i,j})\right]\left(\frac{1}{\alpha}+\sum_{k=1}^{J_i}\frac{R_{i,k}}{J_i^2}\right)-\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{\mu_h}{\alpha}\right\}$$
$$=\frac{1}{N\bar{J}}\left\{\sum_{i=1}^{N}\left[\sum_{j=1}^{J_i}\frac{\tilde{h}(x_{i,j})-\mu_h}{\alpha}\right]+\sum_{i=1}^{N}\left[\sum_{j=1}^{J_i}\tilde{h}(x_{i,j})\right]\left[\sum_{k=1}^{J_i}\frac{R_{i,k}}{J_i^2}\right]\right\},$$

where $R_{i,j}$ is the remainder term in Lemma 5. This converges to zero since $R_{i,k}$ is bounded and $\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\tilde{h}(x_{i,j})\xrightarrow{p}\mu_h$ by the law of large numbers. *Q.E.D.*

LEMMA 7: *Under the conditions of Theorem* 4,

$$\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}x_{i,j}\left(p_{i,j}-MC_{i,j}-\frac{1}{\alpha}\right)=\frac{1}{\bar{J}}\left(\frac{\mu_{xr}}{\alpha\mu_r}+o_P(1)\right).$$

PROOF: We have

$$\bar{J}\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}x_{i,j}\left(p_{i,j}-MC_{i,j}-\frac{1}{\alpha}\right)$$
$$=\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J_i}x_{i,j}\left(\frac{r_{i,j}}{\alpha\sum_{k=1}^{J_i}r_{i,k}}+\frac{R_{i,j}}{J_i^2}\right)$$
$$=\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sum_{j=1}^{J_i}x_{i,j}r_{i,j}}{\alpha\sum_{k=1}^{J_i}r_{i,k}}+\sum_{j=1}^{J_i}\frac{x_{i,j}R_{i,j}}{J_i^2}\right),$$

where $R_{i,j}$ is the quantity in Lemma 5. The last term is bounded by a constant times $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{J_i} \leq \frac{1}{\min_{1\leq i\leq N} J_i}$, which goes to zero under the conditions of Theorem 4. For the first term, we have

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\sum_{j=1}^{J_i} x_{i,j}r_{i,j}}{\alpha\sum_{k=1}^{J_i} r_{i,k}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\sum_{j=1}^{J_i} x_{i,j}r_{i,j}}{\alpha J_i\mu_r}\frac{J_i\mu_r}{\sum_{k=1}^{J_i} r_{i,k}}$$

$$= \frac{\mu_{xr}}{\alpha\mu_r} + \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\frac{1}{J_i}\sum_{j=1}^{J_i} x_{i,j}r_{i,j} - \mu_{xr}}{\alpha\mu_r}\right)$$

$$+ \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\frac{1}{J_i}\sum_{j=1}^{J_i} x_{i,j}r_{i,j}}{\alpha\mu_r}\right)\left(\frac{\mu_r}{\frac{1}{J_i}\sum_{k=1}^{J_i} r_{i,k}} - 1\right).$$

The second term has mean zero and variance $\frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{J_i^2}\operatorname{var}(x_{i,j}r_{i,j}/(\alpha\mu_r))$, which goes to zero as $N \to \infty$, and the last term is equal to

$$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\frac{1}{J_i}\sum_{j=1}^{J_i} x_{i,j}r_{i,j}}{\alpha\mu_r\frac{1}{J_i}\sum_{k=1}^{J_i} r_{i,k}}\right)\left(\mu_r - \frac{1}{J_i}\sum_{k=1}^{J_i} r_{i,k}\right).$$

By boundedness of $x_{i,j}$ and $r_{i,j}$, this is bounded by a constant times $\frac{1}{N}\sum_{i=1}^{N}|\mu_r - \frac{1}{J_i}\sum_{k=1}^{J_i} r_{i,k}|$, which converges in probability to zero since, by Hölder's inequality,

$$\frac{1}{N}\sum_{i=1}^{N} E\left|\mu_r - \frac{1}{J_i}\sum_{k=1}^{J_i} r_{i,k}\right| \leq \frac{1}{N}\sum_{i=1}^{N}\sqrt{E\left(\mu_r - \frac{1}{J_i}\sum_{k=1}^{J_i} r_{i,k}\right)^2}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sqrt{\operatorname{var}(r_{i,k})/J_i} \to 0.$$

Q.E.D.

LEMMA 8: *Under the conditions of Theorem* 4, *for any random variables* $v_{i,j}$ *that are i.i.d. over both* i *and* j *with mean* μ_v *and a finite fourth moment*,

$$\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left[\frac{1}{\bar{J}}\sum_{k\neq j}\tilde{h}(x_{i,k})\right]v_{i,j} \tag{9}$$

$$= \tilde{m}_2\mu_h\mu_v + \frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}\left[\mu_h(v_{i,j}-\mu_v)+\mu_v\big(\tilde{h}(x_{i,j})-\mu_h\big)\right]$$

$$+ o_P(1/\sqrt{N\bar{J}}).$$

PROOF: We have

$$\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left[\frac{1}{\bar{J}}\sum_{k\neq j}\tilde{h}(x_{i,k})\right]v_{i,j} \tag{10}$$

$$= \frac{1}{N\bar{J}^2}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{k\neq j}\big[\mu_h\mu_v + \mu_h(v_{i,j}-\mu_v)+\mu_v\big(\tilde{h}(x_{i,k})-\mu_h\big)$$

$$+ (v_{i,j}-\mu_v)\big(\tilde{h}(x_{i,k})-\mu_h\big)\big]$$

$$= \mu_h\mu_v\frac{1}{N\bar{J}^2}\sum_{i=1}^{N}J_i(J_i-1)$$

$$+ \frac{1}{N\bar{J}^2}\sum_{i=1}^{N}\sum_{j=1}^{J_i}(J_i-1)\big[\mu_h(v_{i,j}-\mu_v)+\mu_v\big(\tilde{h}(x_{i,j})-\mu_h\big)\big]$$

$$+ \frac{1}{N\bar{J}^2}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{k\neq j}(v_{i,j}-\mu_v)\big(\tilde{h}(x_{i,k})-\mu_h\big).$$

The first term is equal to $(m_2-1/\bar{J})\mu_h\mu_v = \tilde{m}_2\mu_h\mu_v$. The second term is equal to the second term in (9) minus $\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{1}{\bar{J}}[\mu_h(v_{i,j}-\mu_v)+\mu_v(\tilde{h}(x_{i,j})-\mu_h)]$, which is $\mathcal{O}_P(1/(\bar{J}\sqrt{N\bar{J}})) = o_P(1/\sqrt{N\bar{J}})$ by the central limit theorem. The last term in (10) has mean zero and variance given by

$$\frac{1}{N^2\bar{J}^4}E\left[\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{k\neq j}(v_{i,j}-\mu_v)\big(\tilde{h}(x_{i,k})-\mu_h\big)\right]^2$$

$$= \frac{1}{N^2\bar{J}^4}\sum_{i=1}^{N}E\left[\sum_{j=1}^{J_i}\sum_{k\neq j}(v_{i,j}-\mu_v)\big(\tilde{h}(x_{i,k})-\mu_h\big)\right]^2$$

$$= \frac{1}{N^2\bar{J}^4}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\sum_{k\neq j}\sum_{\ell=1}^{J_i}\sum_{m\neq \ell} E(v_{i,j}-\mu_v)\big(\tilde{h}(x_{i,k})-\mu_h\big)$$

$$\times (v_{i,\ell}-\mu_v)\big(\tilde{h}(x_{i,m})-\mu_h\big).$$

For each i, all of the terms in the above summand are zero except for those where either $j=\ell$ and $k=m$ or $j=m$ and $k=\ell$. The number of such terms is bounded by a constant times J_i^2, so that the above display is bounded by a constant times $\frac{1}{N^2\bar{J}^2}\sum_{i=1}^{N}(J_i/\bar{J})^2 = \frac{1}{N\bar{J}^2}m_2$. Thus, the last term in (10) converges to zero at a $1/\sqrt{N\bar{J}^2}$ rate, which is strictly faster than $1/\sqrt{N\bar{J}}$, as claimed. *Q.E.D.*

LEMMA 9: *Under the conditions of Theorem* 4,

$$\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left(\begin{array}{c} x_{i,j} \\ \frac{1}{\bar{J}}\sum_{k\neq j}\tilde{h}(x_{i,k})\end{array}\right)\left(x_{i,j}' \quad MC_{i,j}+\frac{1}{\alpha}\right) = \tilde{M}_1 + V_{JN}/\sqrt{N\bar{J}},$$

where $\tilde{M}_1$ *is defined at the beginning of this section and* $(\mathrm{vec}(V_{JN})', (\frac{1}{\sqrt{N\bar{J}}}\sum_{i=1}^{N}\sum_{j=1}^{J_i} z_{i,j}\xi_{i,j})')'$ *converges to a normal distribution with variance* $\widetilde{\Sigma}_{Wu}$ (*where* $\widetilde{\Sigma}_{Wu}$ *is defined at the beginning of this section*).

PROOF: It follows from Lemma 8 that $(\mathrm{vec}(V_{JN})', (\frac{1}{\sqrt{N\bar{J}}}\sum_{i=1}^{N}\sum_{j=1}^{J_i} z_{i,j}\xi_{i,j})')'$ is, up to $o_P(1)$, equal to

$$\frac{1}{\sqrt{N\bar{J}}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}(I_{d_x+2}\otimes B_{J_i/\bar{J}})\begin{pmatrix}\mathrm{vec}(W_{i,j})\\ u_{i,j}\end{pmatrix},$$

where $W_{i,j}$ and $u_{i,j}$ are defined at the beginning of this section and, for any scalar r, B_r is defined to be the $d_z\times d_z$ diagonal matrix with 1's in the first d_x diagonal entries and r in the remaining diagonal entries. By a central limit theorem for triangular arrays of independent nonidentically distributed variables, this converges to a normal distribution with variance

$$\lim_{N\to\infty}\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}(I_{d_x+2}\otimes B_{J_i/\bar{J}})\Sigma_{Wu}(I_{d_x+2}\otimes B_{J_i/\bar{J}})',$$

which can be seen to be equal to $\widetilde{\Sigma}_{Wu}$ by inspection. (To verify Lindeberg's condition for the terms of the form $(J_i/\bar{J})v_{i,j}$ for a random variable $v_{i,j}$, it suffices to show that $\frac{\max_{1\le i\le N}(J_i/\bar{J})^2}{\sum_{i=1}^N\sum_{j=1}^{J_i}(J_i/\bar{J})^2}\to 0$, which follows since $\frac{\max_{1\le i\le N}(J_i/\bar{J})^2}{\sum_{i=1}^N\sum_{j=1}^{J_i}(J_i/\bar{J})^2}\le \frac{\max_{1\le i\le N}(J_i/\bar{J})^2}{\max_{1\le i\le N}\sum_{j=1}^{J_i}(J_i/\bar{J})^2}=\frac{1}{\max_{1\le i\le N}J_i}\to 0$.) *Q.E.D.*

Putting the above lemmas together and using the fact that $\tilde{M}_1=K_{\tilde{m}_2}M_1H$, we have

$$\begin{aligned}\hat{M}_{zx}&\equiv\frac{1}{N\bar{J}}\sum_{i=1}^N\sum_{j=1}^{J_i}z_j\big(x_j',p_j\big)\\&=K_{\tilde{m}_2}M_1H+V_{JN}/\sqrt{\bar{J}N}+M_2\big(A_{1/\bar{J}}+o_P(1/\bar{J})\big),\end{aligned}$$

where V_{JN} is given in Lemma 9. Since the last column of M_1 is all 0's, $M_1A_{\bar{J}}=M_1$. Also, since the first d_x columns of $M_2A_{1/\bar{J}}$ are zero, $M_2A_{1/\bar{J}}H=M_2A_{1/\bar{J}}$, so that $M_2A_{1/\bar{J}}H^{-1}=M_2A_{1/\bar{J}}$ as well. Thus,

$$Q_N\equiv\hat{M}_{zx}H^{-1}A_{1/\bar{J}}^{-1}=K_{\tilde{m}_2}M_1+V_{JN}H^{-1}A_{\bar{J}}/\sqrt{\bar{J}N}+M_2+o_P(1).$$

Let $\hat{Z}_{z\xi}=\frac{1}{\sqrt{\bar{J}N}}\sum_{i=1}^N\sum_{j=1}^{J_i}z_{i,j}\xi_{i,j}$.

It follows that, in the case where $N/\bar{J}\to\infty$,

$$\begin{aligned}&\operatorname{diag}(\sqrt{\bar{J}N},\dots,\sqrt{\bar{J}N},\sqrt{N/\bar{J}})H\big[\big(\hat{\beta}',-\hat{\alpha}\big)'-\big(\beta',-\alpha\big)'\big]\\&\quad=\sqrt{\bar{J}N}A_{1/\bar{J}}H\big[\big(\hat{\beta}',-\hat{\alpha}\big)'-\big(\beta',-\alpha\big)'\big]\\&\quad=\big(Q_N'WQ_N\big)^{-1}Q_N'W\hat{Z}_{z\xi}\xrightarrow{d}\big(Q'WQ\big)^{-1}Q'WZ_{z\xi},\end{aligned}$$

where Q is the deterministic matrix and $Z_{z\xi}$ is the random vector defined at the beginning of the proof. In the case where $N/\bar{J}\to c$ for a finite constant c,

$$\begin{aligned}\tilde{Q}_N&\equiv\hat{M}_{zx}H^{-1}A_{1/\sqrt{N\bar{J}}}^{-1}\\&=K_{m_2}M_1+V_{JN}H^{-1}A_{\sqrt{N\bar{J}}}/\sqrt{\bar{J}N}+M_2A_{\sqrt{N/\bar{J}}}+o_P(1).\end{aligned}$$

This converges in distribution to $K_{m_2,\infty}M_1+Z_{zx}H^{-1}\operatorname{diag}(0,\dots,0,1)+M_2A_{\sqrt{c}}=\tilde{Q}_{\infty,c}$ jointly with $\hat{Z}_{z\xi}$. (Here, $(\operatorname{vec}(Z_{zx})',Z_{z\xi}')'$ is normal with mean zero and variance matrix $\widetilde{\Sigma}_{Wu}$ as defined at the beginning of the proof. Note that $\hat{Z}_{z\xi}$ and V_{JN} converge in distribution jointly to $Z_{z\xi}$ and Z_{zx} by Lemma 9.) Thus,

$$\operatorname{diag}(\sqrt{\bar{J}N},\dots,\sqrt{\bar{J}N},1)H\big[\big(\hat{\beta}',-\hat{\alpha}\big)'-\big(\beta',-\alpha\big)'\big]$$

$$= \sqrt{\bar{J}N} A_{1/\sqrt{NJ}} H\big[(\hat{\beta}', -\hat{\alpha})' - (\beta', -\alpha)'\big]$$

$$= (\tilde{Q}_N' W \tilde{Q}_N)^{-1} \tilde{Q}_N' W \hat{Z}_{z\xi} \xrightarrow{d} (\tilde{Q}_{\infty,c}' W \tilde{Q}_{\infty,c})^{-1} \tilde{Q}_{\infty,c}' W Z_{z\xi}.$$

For $c = 0$, $\tilde{Q}_{\infty,c} = K_{m_2,\infty} M_1 + Z_{zx} H^{-1} \operatorname{diag}(0, \dots, 0, 1)$, and this limiting distribution is the same as if the markup were equal to $1/\alpha$ (by the same arguments, but with M_2 a matrix of zeros).

B. ADDITIONAL LARGE MARKET ASYMPTOTIC RESULTS

This section gives the formal results described in Section 3.2 for the nested logit model, and discusses large market asymptotics for the vertical model, and for some of the cases considered in the main text under multi-product firms.

B.1. *Nested Logit*

In the nested logit model, the J products are split into G mutually exclusive groups. Here, the number of groups G will increase, while the number of products per group stays fixed. As in Section 3.1, this section considers single product firms, although the results will be similar for multi-product firms as long as the number of firms increases rather than the number of products per firm. The set of products in a given group $g \in \{1, \dots, G\}$ is denoted by $\mathcal{J}_g \subseteq \{1, \dots, J\}$. The share of product j as a fraction of its group g is denoted by $\bar{s}_{j/g}(x, p, \xi)$, and the share of group g as a fraction of all products is given by $\bar{s}_g(x, p, \xi)$. Consumer i's utility for good j is

$$u_{ij} = x_j'\beta - \alpha p_j + \xi_j + \zeta_{ig} + (1 - \sigma)\varepsilon_{ij} \equiv \delta_j + \zeta_{ig} + (1 - \sigma)\varepsilon_{ij},$$

where ζ_{ig} is a random coefficient on a dummy variable for group g and ε_{ij} is still extreme value. The distribution of ζ_{ig} depends on σ and is such that $\zeta_{ig} + (1 - \sigma)\varepsilon_{ij}$ is extreme value. This leads to the formulas $\bar{s}_{j/g} = \frac{\exp(\delta_j/(1-\sigma))}{D_g}$ and $\bar{s}_g = \frac{D_g^{1-\sigma}}{\sum_h D_h^{1-\sigma}}$ for shares where $D_g = \sum_{j \in \mathcal{J}_g} \exp(\delta_j/(1 - \sigma))$. These can be inverted to get

$$\log s_j - \log s_0 = x_j'\beta - \alpha p_j + \sigma \log \bar{s}_{j/g} + \xi_j \tag{11}$$

(here, the outside good, product 0, has mean utility normalized to zero and is the only product in its nest). The derivative of j's share with respect to j's price is $\frac{ds_j}{dp_j} = \frac{-\alpha}{1-\sigma} s_j(1 - \sigma \bar{s}_{j/g} - (1 - \sigma)s_j)$, which gives a markup of

$$p_j - MC_j = \frac{1 - \sigma}{\alpha} \Big/ \big(1 - \sigma \bar{s}_{j/g} - (1 - \sigma)s_j\big). \tag{12}$$

If the number of nests increases with the number of products per nest fixed, s_j will go to zero. Thus, we might expect that prices converge to the solution to a limiting system of equations where s_j is removed from the right hand side of (12). Since $\bar{s}_{j/g}$ depends only on products in group g, this would mean that asymptotic markups are determined by a pricing game involving only firms with products in the same group. To formalize this, let p_j^* for j in group j be defined as the unique solution to the system of equations

$$
\begin{aligned}
(13)\qquad p_j^* - MC_j &= \frac{1-\sigma}{\alpha} \Big/ \big(1 - \sigma \bar{s}_{j/g}(x, p^*, \xi)\big) \\
&= \frac{1-\sigma}{\alpha} \Bigg(\sum_{k \in \mathcal{J}_g} \exp\big((x_k'\beta - p_k^*\alpha + \xi_k)/(1-\sigma)\big) \Bigg) \\
&\quad \Bigg/ \Bigg(\Bigg[\sum_{k \in \mathcal{J}_g} \exp\big((x_k'\beta - p_k^*\alpha + \xi_k)/(1-\sigma)\big) \Bigg] \\
&\quad - \sigma \exp\big((x_j'\beta - p_j^*\alpha + \xi_j)/(1-\sigma)\big) \Bigg)
\end{aligned}
$$

and let $\bar{s}_{j/g}^* = \bar{s}_{j/g}(x, p^*, \xi)$. That is, p_j^* is defined as the solution to a system of equations given by the markup formula (12), but with s_j set to its limiting value of 0. The following theorem states that IV estimates in this model are asymptotically equivalent to the estimates that would be obtained if prices were replaced with p_j^*. Since prices in the limiting model depend on characteristics of products in the same nest but not on characteristics of products in other nests, this means that characteristics of products in the same nest will potentially have identifying power, while products in other nests will not.

THEOREM 7: *In the nested logit model single product firms and many nests, suppose that* (x_j, ξ_j, MC_j) *is bounded and i.i.d. across* j. *Let* $z_j = (x_j, h(\{x_k\}_{k\in\mathcal{J}_{g-L}}, \ldots, \{x_k\}_{k\in\mathcal{J}_{g+M}}))$ *for* $j \in \mathcal{J}_g$ *for some function* h *with finite variance. Let* p_j^* *and* $\bar{s}_{j/g}^*$ *be defined in* (13). *Let* $(\hat{\alpha}, \hat{\beta}, \hat{\sigma})$ *be the IV estimates defined in* (4), *and let* $(\hat{\alpha}^*, \hat{\beta}^*, \hat{\sigma}^*)$ *be defined in the same way, but with* p_j^* *and* $\bar{s}_{j/g}^*$ *replacing* p_j *and* $\bar{s}_{j/g}$. *Then* $\|(\hat{\alpha}, \hat{\beta}, \hat{\sigma}) - (\hat{\alpha}^*, \hat{\beta}^*, \hat{\sigma}^*)\| \xrightarrow{p} 0$ *and, if* $(\hat{\alpha}^*, \hat{\beta}^*, \hat{\sigma}^*)$ *is consistent and asymptotically normal*, $(\hat{\alpha}, \hat{\beta}, \hat{\sigma})$ *will also be consistent and asymptotically normal, with the same asymptotic distribution.*

Note that if we had taken the number of nests fixed with the number of products per nest increasing, both $\bar{s}_{j/g}$ and s_j would converge to zero in the markup formula (12), and the markup would converge to a constant as with the results in Section 3.1. Thus, if the dimension of ζ is fixed, we obtain the same results as in Section 3.1 (with the stronger result for the nested logit model that $\|\hat{\sigma}-\hat{\sigma}^*\| \xrightarrow{p} 0$, where both estimates are inconsistent). The proof of Theorem 7 is given below.

PROOF OF THEOREM 7: As before, it suffices to show that p_j and, in this case, $s_{j/g}$ converge uniformly to the starred versions at a faster than $1/\sqrt{J}$ rate.

Define the function f by

$$\begin{aligned}
&f_j(p,x,\xi,MC,r)\\
&\quad = p_j - MC_j\\
&\qquad - \frac{1-\sigma}{\alpha}\Bigl(\sum_k \exp\bigl((x_k'\beta - p_k\alpha+\xi_k)/(1-\sigma)\bigr)\Bigr)\\
&\qquad \Big/\biggl(\Bigl[\sum_k \exp\bigl((x_k'\beta - p_k\alpha+\xi_k)/(1-\sigma)\bigr)\Bigr]\\
&\qquad - \sigma\exp\bigl((x_j'\beta - p_j\alpha+\xi_j)/(1-\sigma)\bigr)\biggr) + r_j.
\end{aligned}$$

Then p_g^* satisfies $f(p_g^*, x_g, \xi_g, MC_g, 0) = 0$ and any solution p to the Nash pricing equations satisfies $f(p_g^*, x_g, \xi_g, MC_g, \tilde{r}) = 0$ for

$$\tilde{r}_j = \frac{1-\sigma}{\alpha}\frac{(1-\sigma)s_j(p,x)}{\bigl(1-\sigma\bar{s}_{j/g}(p,x)\bigr)\bigl(1-\sigma\bar{s}_{j/g}(p,x)-(1-\sigma)s_j(p,x)\bigr)},$$

where the functions s_j and $\bar{s}_{j/g}$ take prices and product characteristics to the expressions for nested logit shares defined earlier in the section.

The proof proceeds by first showing that $\sqrt{J}\max_{j\le J}\tilde{r}_j$ converges to zero, and then using the implicit function theorem and the mean value theorem to get a linear approximation to the p that solves $f(p,x,\xi,MC,r)=0$ as a function of r. The first statement follows since

$$|\tilde{r}_j| \le \frac{1-\sigma}{\alpha}\frac{s_j(p,x)}{1-\sigma-(1-\sigma)s_j(p,x)},$$

so that $\sqrt{J}\max_{j\le J}\tilde{r}_j$ will converge to zero as long as $\sqrt{J}\max_{j\le J}s_j$ converges to zero. Inspection of the formula for s_j shows that this will hold as long as equilibrium prices are bounded.

For r small, the equation $f(p,x,\xi,MC,r)=0$ has a unique solution for p. To see that a solution exists, note that this equation is equivalent to the first order condition for setting prices in the Bertrand pricing game with demand given by $q_j(p) \equiv \exp((x_j'\beta - \alpha p_j)/(1-\sigma))/D_g^{\sigma}$ and marginal cost equal to $MC_j + r_j$. An equilibrium exists in this game, since it is log supermodular (see pp. 151–152 of Vives (2001)):

$$\begin{aligned}
\frac{\partial^2 \log \pi_j}{\partial p_j\,\partial p_k} &= \frac{\partial^2 \log q_j(p)}{\partial p_j\,\partial p_k} \\
&= \frac{\partial^2}{\partial p_j\,\partial p_k}\Big\{\log \exp\big((x_j'\beta - \alpha p_j)/(1-\sigma)\big) \\
&\qquad - \sigma \log \sum_\ell \exp\big((x_\ell'\beta - \alpha p_\ell)/(1-\sigma)\big)\Big\} \\
&= -\frac{\partial}{\partial p_j}\sigma \frac{\frac{-\alpha}{1-\sigma}\exp\big((x_k'\beta - \alpha p_k)/(1-\sigma)\big)}{\sum_\ell \exp\big((x_\ell'\beta - \alpha p_\ell)/(1-\sigma)\big)} \\
&= \frac{\alpha\sigma \exp\big((x_k'\beta - \alpha p_k)/(1-\sigma)\big)}{1-\sigma} \\
&\qquad \times \frac{\frac{\alpha}{1-\sigma}\exp\big((x_j'\beta - \alpha p_j)/(1-\sigma)\big)}{\Big(\sum_\ell \exp\big((x_\ell'\beta - \alpha p_\ell)/(1-\sigma)\big)\Big)^2} \\
&> 0.
\end{aligned}$$

Uniqueness follows from verifying a dominant diagonal condition for f (see p. 47 of Vives (2001)). We have

$$\begin{aligned}
\frac{\partial f_j}{\partial p_j} &= 1 - \frac{1-\sigma}{\alpha}\sigma \frac{1}{\big(1-\sigma \bar{s}_{j/g}(p)\big)^2}\frac{\partial}{\partial p_j}\bar{s}_{j/g}(p) \\
&= 1 - \frac{1-\sigma}{\alpha}\sigma \frac{1}{\big(1-\sigma \bar{s}_{j/g}(p)\big)^2}\frac{-\alpha}{1-\sigma}\bar{s}_{j/g}(p)\big(1-\bar{s}_{j/g}(p)\big) \\
&= 1 + \sigma \frac{\bar{s}_{j/g}(p)\big(1-\bar{s}_{j/g}(p)\big)}{\big(1-\sigma \bar{s}_{j/g}(p)\big)^2}
\end{aligned}$$

and, for $k \neq j$,

$$\begin{aligned}
\frac{\partial f_j}{\partial p_k} &= -\frac{1-\sigma}{\alpha}\sigma \frac{1}{\big(1-\sigma \bar{s}_{j/g}(p)\big)^2}\frac{\partial}{\partial p_k}\bar{s}_{j/g}(p) \\
&= -\frac{1-\sigma}{\alpha}\sigma \frac{1}{\big(1-\sigma \bar{s}_{j/g}(p)\big)^2}\frac{\alpha}{1-\sigma}\bar{s}_{j/g}(p)\bar{s}_{k/g}(p) \\
&= -\sigma \frac{\bar{s}_{j/g}(p)\bar{s}_{k/g}(p)}{\big(1-\sigma \bar{s}_{j/g}(p)\big)^2}.
\end{aligned}$$

Thus,

$$\frac{\partial f_j}{\partial p_j} - \sum_{k \neq j} \left| \frac{\partial f_j}{\partial p_k} \right| = 1 + \frac{\sigma \bar{s}_{j/g}(p)}{\left(1 - \sigma \bar{s}_{j/g}(p)\right)^2} \left(1 - s_{j/g}(p) - \sum_{k \neq j} s_{k/g}(p)\right)$$
$$= 1 > 0.$$

Since a unique p solves $f(p, x, \xi, MC, r) = 0$ for the elements of (x, ξ, MC) in the given bounded set and r close to zero, this defines p as a function $\phi(x, \xi, MC, r)$ of the remaining variables. By the implicit function theorem, the derivative matrix $D\phi(x, \xi, MC, r)$ of ϕ is given by

$$\left(D_p f\left(\phi(x, \xi, MC, r), x, \xi, MC, r\right)\right)^{-1}$$
$$\times D_{x,\xi,MC,r} f\left(\phi(x, \xi, MC, r), x, \xi, MC, r\right),$$

where subscripts denote blocks of the derivative matrix corresponding to derivatives with respect to given variables (the derivative matrix of f with respect to p is invertible since it is diagonally dominant). Since $p = \phi(x, \xi, MC, \tilde{r})$ and $p^* = \phi(x, \xi, MC, 0)$, by the mean value theorem, for every index j, there is a $\bar{r}$ between 0 and $\tilde{r}$ such the difference between p_j and p_j^* is given by the jth row of

$$\left(D_p f\left(\phi(x, \xi, MC, \bar{r}), x, \xi, MC, \bar{r}\right)\right)^{-1}$$
$$\times D_r f\left(\phi(x, \xi, MC, \bar{r}), x, \xi, MC, \bar{r}\right)\tilde{r}.$$

Since the elements of $(D_p f(\phi(x, \xi, MC, r), x, \xi, MC, r))^{-1} D_r f(\phi(x, \xi, MC, r), x, \xi, MC, r)$ are continuous functions of x, ξ, MC, and r, the function that maps t to the maximum of the absolute values of the elements of this matrix times t takes a maximum M as x, ξ, MC, and r range over the compact set that contains them and t ranges over the unit sphere in $\mathbb{R}^{|\mathcal{J}_g|}$. This gives

$$\sqrt{J} \max_{j \leq J} \left| p_j^* - p_j \right| \leq \sqrt{J} \max_{j \leq J} M \|\tilde{r}_j\| \to 0.$$

The rate of uniform convergence for $\bar{s}_{j/g}$ follows since $\bar{s}_{j/g}$ is equal to $\bar{s}_{j/g}^*$ with p_k^* replaced by p_k in the definition, and the formula in the definition has a derivative with respect to the vector of prices in group g that is bounded in an open set containing all values of (x, ξ, MC, p) that can be taken under the assumptions of the theorem. Thus, by the mean value theorem, for some finite B, $\sqrt{J} \max_{j \leq J} |\bar{s}_{j/g}^* - \bar{s}_{j/g}| \leq \sqrt{J} B \max_{j \leq J} |p_j^* - p_j| \to 0.$ *Q.E.D.*

B.2. *Vertical Model*

In contrast to the other models in which consumers have an idiosyncratic preference term ε_{ij} for each item, consider a model in which consumers agree on the ranking of goods, but differ in their willingness to pay for product quality, as in Bresnahan (1987). As with the random coefficients logit results in Section 3.1, the identifying power of characteristics of other products as instruments for price goes to zero at a faster than $1/\sqrt{J}$ rate.[13]

Utility of an individual consumer is given by

$$u_{ij} = x_j'\beta - \zeta_{ip} p_j + \xi_j \equiv \delta_j - \zeta_{ip} p_j,$$

where ζ_{ip} represents consumer i's preference for product quality. A small value of ζ_{ip} means that consumer i has a high value for the quality of the inside goods relative to the numeraire good. The outsize good 0 has $p_0 = 0$ and δ_0 normalized to 0.

Arrange the goods in order of product quality so that $\delta_1 < \cdots < \delta_J$. If all products have positive market share, this will imply that prices satisfy $p_1 < \cdots < p_J$ as well. Consumer i will prefer good j to $j-1$ if

$$\delta_j - \zeta_{ip} p_j > \delta_{j-1} - \zeta_{ip} p_{j-1} \quad \Longleftrightarrow \quad \Delta_j \equiv \frac{\delta_j - \delta_{j-1}}{p_j - p_{j-1}} > \zeta_{ip}.$$

Combining this with the expression for $j+1$, consumer i will prefer j to its neighbors if $\Delta_j > \zeta_{ip} > \Delta_{j+1}$. For all products to have positive market share, this must hold for some ζ_{ip} for all j, so we must have $\Delta_1 > \cdots > \Delta_J$. If this is the case, consumers who prefer j to its neighbors will also prefer j to all other products, so, letting F be the cumulative distribution function (c.d.f.) of ζ_{ip}, market shares will be given by

$$s_j = F(\Delta_j) - F(\Delta_{j+1}). \tag{14}$$

If we define $\Delta_0 = \infty$ and $\Delta_{J+1} = -\infty$, this will hold for good J and the outside good 0 as well.

This can be inverted to give

$$F^{-1}\left(\sum_{k=j}^{J} s_k\right)(p_j - p_{j-1}) = (x_j - x_{j-1})'\beta + \xi_j - \xi_{j-1}. \tag{15}$$

If F is known, this equation can be estimated using OLS (indeed, Bresnahan (1987), treats F as known and fixes F when estimating a version of this model). If F is allowed to depend on an unknown parameter (as in, e.g., Berry and Pakes (2007)), more instruments will be needed, so it will be useful to study the identifying power of moment conditions based on characteristics of other

[13] Note, however, that the version of this model used by Bresnahan (1987) places enough structure on the distribution of random coefficients that the model is identified through other means (see the discussion surrounding equation (15) below).

products (note, however, that, unless the parameter enters linearly into (15), the inconsistency results in this paper will not apply, and additional arguments will be needed).

Differentiating the formula for shares with respect to p_j gives, letting f be the probability density function (p.d.f.) of ζ_{ip},

$$\frac{ds_j}{dp_j} = -f(\Delta_j)\frac{\Delta_j}{p_j - p_{j-1}} - f(\Delta_{j+1})\frac{\Delta_{j+1}}{p_{j+1} - p_j}.$$

This gives markups in an interior Bertrand equilibrium as

$$p_j - MC_j = \frac{F(\Delta_j) - F(\Delta_{j+1})}{f(\Delta_j)\dfrac{\Delta_j}{p_j - p_{j-1}} + f(\Delta_{j+1})\dfrac{\Delta_{j+1}}{p_{j+1} - p_j}}. \tag{16}$$

Suppose that, for some $\underline{\zeta} > 0$, $\underline{\zeta} \le \zeta_{ip}$ for all consumers. That is, willingness to pay for product quality is bounded from above. In this case, if all products have positive market share, we will have $\Delta_j > \underline{\zeta}$ for all j. Thus, the denominator in (16) will be bounded from below as J increases, so if market shares all converge to zero, markups will converge to zero at the same rate or faster. If firms have approximately equal market shares asymptotically, they will converge to zero at a $1/J$ rate, fast enough for Theorem 5 to hold.

One set of primitive conditions under which markups will converge to zero at a fast rate is the following. In addition to assuming that ζ_{ip} is bounded from below, suppose that the density f of the random coefficient is bounded from above by $\overline{f}$ and from below by $\underline{f}$. Suppose that product characteristics are added in such a way that $\sqrt{J}\max_{j\le J}\delta_j - \delta_{j-1} \to 0$ (e.g., this holds with probability 1 by results in Devroye (1981), for the case where the δ_j's are order statistics of the uniform distribution or, by a quantile transformation, any distribution with finite support and continuous density bounded from above and below) and that all products have positive market share in equilibrium. Then

$$\begin{aligned}
p_j - MC_j &= \frac{F(\Delta_j) - F(\Delta_{j+1})}{f(\Delta_j)\dfrac{\Delta_j}{p_j - p_{j-1}} + f(\Delta_{j+1})\dfrac{\Delta_{j+1}}{p_{j+1} - p_j}} \\
&\le \frac{\overline{f}}{\underline{f}}\frac{\Delta_j - \Delta_{j+1}}{\dfrac{\Delta_j}{p_j - p_{j-1}} + \dfrac{\Delta_{j+1}}{p_{j+1} - p_j}} \le \frac{\overline{f}}{\underline{f}}(p_j - p_{j-1})
\end{aligned}$$

(the last inequality follows by bounding the denominator from below by $\underline{f}\frac{\Delta_j - \Delta_{j+1}}{p_j - p_{j-1}}$). For product j to have positive market share, we must have

$$\underline{\zeta} < \frac{\delta_j - \delta_{j-1}}{p_j - p_{j-1}} \quad \Longrightarrow \quad p_j - p_{j-1} < \frac{\delta_j - \delta_{j-1}}{\underline{\zeta}}.$$

Thus,

$$\sqrt{J} \max_{j \le J} (p_j - MC_j) \le \sqrt{J} \frac{\overline{f}}{\underline{f} \cdot \underline{\zeta}} \max_{j \le J} (\delta_j - \delta_{j-1}) \to 0.$$

We note that, while the above conditions lead to markups quickly decreasing, the results may be different if the support of product characteristics or the distribution of the random coefficient changes with J. We leave these questions for future research.

B.3. *Multi-Product Firms*

This section considers the case with many small multi-product firms. If the number of products sold by each firm is fixed and the number of firms grows large, the results are similar to the single product case, although, due to the difficulty of proving existence and uniqueness of equilibrium for these models with multi-product firms, these results place some conditions directly on equilibrium prices. In particular, these results require prices to be bounded as the number of products increases, and the nested logit model requires the existence of an equilibrium in a limiting form of the game in which price is a differentiable function of costs and characteristics.

For the logit model, we have $\frac{\partial s_j}{\partial p_j} = -\alpha s_j(1 - s_j)$ and, for $k \neq j$, $\frac{\partial s_j}{\partial p_k} = \alpha s_j s_k$. Substituting this into the first order conditions for p_j (equation (1)) and dividing by $-\alpha s_j$ gives

$$(17) \qquad (p_j - MC_j)\big(1 - s_j(x, p, \xi)\big) - \sum_{k \in \mathcal{F}_f, k \neq j} (p_k - MC_k) s_k(x, p, \xi) - \frac{1}{\alpha} = 0.$$

Assuming that prices and product characteristics are bounded as J increases, shares will go to zero at a faster than $1/\sqrt{J}$ rate. In this case, markups will converge to $1/\alpha$ at a faster than $1/\sqrt{J}$ rate, as in the single product case.

For the nested logit model, it can be checked that, for $k \neq j$ and k and j in the same nest, $\partial s_k / \partial p_j = \frac{\alpha}{1-\sigma} s_k(\sigma \bar{s}_{j/g} + (1 - \sigma) s_j)$. For k in some other nest ℓ, we have $\partial s_k/\partial p_j = \alpha s_k s_j$. Plugging these into the first order conditions for firm f setting p_j gives

$$\begin{aligned} 0 = &-\frac{\alpha}{1-\sigma}(p_j - MC_j)s_j\big(1 - \sigma\bar{s}_{j/g} - (1-\sigma)s_j\big) \\ &+ \sum_{k \in \mathcal{F}_f \cap \mathcal{J}_g, k \neq j} (p_k - MC_k)\frac{\alpha}{1-\sigma}s_k\big(\sigma\bar{s}_{j/g} + (1-\sigma)s_j\big) \\ &+ \sum_{k \in \mathcal{F}_f \setminus \mathcal{J}_g} (p_k - MC_k)\alpha s_k s_j + s_j. \end{aligned}$$

Rearranging gives

$$\begin{aligned} 0 = &\frac{1-\sigma}{\alpha} - (p_j - MC_j)\big(1 - \sigma\bar{s}_{j/g} - (1-\sigma)s_j\big) \\ &+ \sum_{k \in \mathcal{F}_f \cap \mathcal{J}_g, k \neq j} (p_k - MC_k)\frac{\bar{s}_{k/g}}{\bar{s}_{j/g}}\big(\sigma\bar{s}_{j/g} + (1-\sigma)s_j\big) \\ &+ \sum_{k \in \mathcal{F}_f \setminus \mathcal{J}_g} (p_k - MC_k)(1-\sigma)s_k. \end{aligned}$$

This can be written—for $\tilde{r}_J$ a term that converges to zero at faster than a $1/\sqrt{J}$ rate as long as prices and product characteristics are bounded as J increases—as

$$\begin{aligned} 0 = &\frac{1-\sigma}{\alpha} - (p_j - MC_j)(1 - \sigma\bar{s}_{j/g}) \\ &+ \sum_{k \in \mathcal{F}_f \cap \mathcal{J}_g, k \neq j} (p_k - MC_k)\sigma\bar{s}_{k/g} + \tilde{r}_J. \end{aligned} \tag{18}$$

If this system of equations has a unique solution, and the function that takes marginal costs and product characteristics of nest g and the remainder term to the vector of prices for nest g that solves this system of equations for nest g has an invertible derivative for marginal costs and product characteristics in a compact set that contains them by assumption, then an argument similar to that used for Theorem 7 will show that prices in the nested logit game converge uniformly at a faster than $1/\sqrt{J}$ rate to those that solve these equations. As with the single product firm case, equilibrium prices do not depend on characteristics of goods in other nests asymptotically. This holds even for products in other nests owned by the same firm.

In the full random coefficients model with multi-product firms, the first order conditions for product j are

$$-\alpha(p_j - MC_j)\int \tilde{s}_j(\delta,\zeta)\big(1-\tilde{s}_j(\delta,\zeta)\big)\,dP_\zeta(\zeta)$$
$$+\alpha\sum_{k\in\mathcal{F}_j,k\neq j}(p_k - MC_k)\int \tilde{s}_j(\delta,\zeta)\tilde{s}_k(\delta,\zeta)\,dP_\zeta(\zeta) + s_j = 0.$$

This can be rearranged to give

$$(p_j - MC_j)\frac{\int \tilde{s}_j(\delta,\zeta)\big(1-\tilde{s}_j(\delta,\zeta)\big)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}$$
$$= \sum_{k\in\mathcal{F}_j,k\neq j}(p_k - MC_k)\frac{\int \tilde{s}_j(\delta,\zeta)\tilde{s}_k(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)} + \frac{1}{\alpha}.$$

If prices are bounded and the assumptions of Theorem 2 hold, the left hand side will converge to $(p_j - MC_j)$ at faster than a $1/\sqrt{J}$ rate. Assuming prices are bounded, the first term on the right hand side is bounded by a constant times $\frac{\int \tilde{s}_j(\delta,\zeta)\tilde{s}_k(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}$. This term goes to zero at the required rate using the same argument as for $\frac{\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}$.

C. MONTE CARLO

This section reports additional details and summary statistics for the Monte Carlos, as well as results for designs not reported in the main text. These results include a comparison to the case where markups are taken to be constant, which gives an idea of how well the conclusion of Theorem 1 regarding large J asymptotics describes the given combinations of N and J for these data generating processes (see Section C.3).

C.1. *Details for the Monte Carlo Designs*

For the Monte Carlos with more than one market, the BLP instruments are formed by taking the excluded instruments for product j in market i, produced by firm f, to be $\sum_{k\in\mathcal{F}_f} x_{i,k}$ and $\sum_{k=1}^{J_i} x_{i,k}$. For the Monte Carlos with BLP instruments in a single market, $\sum_{k=1}^{J_i} x_{i,k}$ is constant, so the excluded instruments are formed as $\sum_{k\in\mathcal{F}_f} x_{i,k}$ and $(\sum_{k\in\mathcal{F}_f} x_{i,k})^2$. For the Monte Carlos with cost shifter instruments, the excluded instruments are $z_{i,j}$ and $z_{i,j}^2$. For the Monte Carlos where prices are generated from constant markups, the form of the instruments is the same (in particular, one of the BLP instruments is still $\sum_{k\in\mathcal{F}_f} x_{i,k}$ even though the ownership structure that defines $\mathcal{F}_f$ does not affect prices). Note that, in all cases, the number of moment conditions is equal to the number of parameters, so the estimator does not depend on the form of the weighting matrix W. All of the Monte Carlo results use 1000 Monte Carlo replications. For a small number of Monte Carlo draws, the equation solver did not converge to a solution for equilibrium prices or the estimator did not converge, and these were discarded.

The share function and inverse share function were computed by Monte Carlo integration with 10 draws of the random coefficients, with the same draws used to generate shares and to compute the inverse share function. Since the same Monte Carlo draws are used in both cases, there is no simulation error from Monte Carlo integration if we consider the random coefficients to be drawn from a discrete distribution with 10 points.

The last two columns report rejection probabilities for a two-sided test for the price coefficient α at its true value and for testing $\alpha = 0$. Note that the second to last column, which gives the rejection probability at the true value of α, is a lower bound for the size of the test, since the size of the test is the supremum of this rejection probability over all possible values of other parameters (correlation between cost shocks and demand shocks, etc.).

C.2. *Additional Summary Statistics and Monte Carlo Designs*

In addition to the Monte Carlos with 10 products per firm, I perform Monte Carlos with 2 products per firm, and with firm size varying between 2 products in approximately 1/3 of the markets, 5 products in 1/3 of the markets, and 10 products per firm in the remaining markets. More precisely, the number of products per firm and the number of products per market for the cases where one or both of these is varied is given as follows. For the cases with 3 markets and the number of products per market varied, the vector of market sizes is $(20, 60, 100)$. For the cases with 20 markets and the number of products per market varied, 7 markets have 20 products, another 7 markets have 60 products, and the remaining 6 markets have 100 products. For the case with 3 markets where the number of products per firm varies, the vector of firm sizes is $(2, 5, 10)$. For the case with 20 markets and firm size varied, 7 markets have 2 products per firm, another 7 have 5 products per firm, and the remaining 6 have 10 products per firm. For the case with 3 markets where both products per market and products per firm vary, one market has 20 products with 5 products per firm, the second market has 60 products with 10 products per firm, and the remaining market has 100 products and 2 products per firm. For the case with 20 markets where both products per market and products per firm vary, 4 markets have 20 products with 2 products per firm, 3 markets have 20 products with 5 products per firm, 4 markets have 60 products with 5 products per firm, 3 markets have 60 products with 10 products per firm, 3 markets have 100 products with 10 products per firm, and the remaining 3 markets have 100 products with 2 products per firm.

These results are reported in Tables IV and V. These tables contain the results from the designs in the main text as well. In addition to the statistics reported in the main text, I also report mean bias and mean absolute deviation from the true value (as opposed to median bias and median absolute deviation, which are reported here as well). Since these estimators are known not to have first moments in similar settings, it may be the case that these quantities are undefined for some of these designs. This may explain some of the erratic

TABLE IV

MONTE CARLO RESULTS FOR BLP INSTRUMENTS

Markets	Firm Size	Products per Market	Median Bias	Median Abs. Dev. From α_0	Mean Bias	Mean Abs. Dev. From α_0	Rejection Prob. at True α	Power of Test of $\alpha = 0$
1	2	20	−0.3385	0.6081	−0.1710	1.0412	0.1439	0.2052
1	2	60	−0.3613	0.6660	−0.2992	1.3802	0.0631	0.0731
1	2	100	−0.3491	0.6825	−0.3345	1.4174	0.1266	0.1628
1	10	20	−0.2147	1.9530	−57.5606	182.1044	0.2729	0.2729
1	10	60	−0.3698	0.6691	−0.1955	1.2607	0.0783	0.1004
1	10	100	−0.3648	0.7177	−0.0373	1.4195	0.1211	0.1381
3	Varied	20	−0.0229	0.1665	0.0392	0.2642	0.0450	0.7390
3	Varied	Varied	−0.0890	0.2786	0.0033	0.4777	0.0520	0.4700
3	Varied	60	−0.0804	0.3922	0.0218	0.7237	0.1002	0.2956
3	Varied	100	−0.1586	0.4504	−0.1198	0.8946	0.0160	0.1590
3	2	20	−0.2893	0.6742	−0.2255	1.6845	0.0280	0.0750
3	2	Varied	−0.3313	0.6753	−0.1899	1.2031	0.0250	0.0600
3	2	60	−0.3697	0.7407	−0.3989	1.4161	0.0090	0.0530
3	2	100	−0.3154	0.7171	−0.2893	1.6900	0.0140	0.0600
3	10	20	−0.1053	0.3358	−0.0006	0.7356	0.0390	0.3980
3	10	Varied	−0.0494	0.2966	0.0890	0.4682	0.1523	0.4649
3	10	60	−0.2186	0.5827	0.1410	1.5941	0.0200	0.1040
3	10	100	−0.2525	0.6383	−0.1924	1.4761	0.1351	0.1762
20	Varied	20	−0.0044	0.0504	−0.0006	0.0614	0.0510	1.0000
20	Varied	Varied	−0.0211	0.1537	0.0031	0.2073	0.0480	0.9170
20	Varied	60	−0.0061	0.1158	0.0066	0.1451	0.0400	0.9990
20	Varied	100	−0.0190	0.1659	0.0136	0.2172	0.0410	0.9450
20	2	20	−0.0393	0.3504	0.0042	1.0057	0.1552	0.4535
20	2	Varied	−0.1578	0.4697	0.0671	1.0065	0.0851	0.2543
20	2	60	−0.1689	0.6458	−0.0580	1.8815	0.0090	0.1080
20	2	100	−0.2191	0.6897	−0.1581	1.7837	0.1061	0.1632
20	10	20	0.0039	0.1140	0.0298	0.1510	0.0390	0.9880
20	10	Varied	−0.0014	0.1001	0.0123	0.1266	0.0400	0.9960
20	10	60	0.0130	0.2345	0.1111	0.4021	0.0230	0.7710
20	10	100	−0.0379	0.3154	0.1358	0.9300	0.0200	0.4560

TABLE V

MONTE CARLO RESULTS FOR COST INSTRUMENTS

Markets	Firm Size	Products per Market	Median Bias	Median Abs. Dev. From α_0	Mean Bias	Mean Abs. Dev. From α_0	Rejection Prob. at True α	Power of Test of $\alpha = 0$
1	2	20	−0.0795	0.3155	0.0105	0.5596	0.1510	0.4387
1	2	60	−0.0202	0.1580	−0.0070	0.2706	0.0893	0.7222
1	2	100	−0.0194	0.1250	−0.0063	0.1916	0.0836	0.7462
1	10	20	−0.0854	0.3049	−0.0639	0.5184	0.0794	0.2487
1	10	60	−0.0247	0.1749	−0.0085	0.3047	0.1130	0.6710
1	10	100	−0.0196	0.1358	−0.0067	0.1980	0.0762	0.7623
3	Varied	20	−0.0241	0.1819	0.0087	0.2797	0.0801	0.6286
3	Varied	Varied	−0.0047	0.0932	0.0059	0.1435	0.0441	0.7854
3	Varied	60	−0.0090	0.0960	0.0077	0.1450	0.0513	0.7678
3	Varied	100	−0.0027	0.0760	0.0123	0.1050	0.0562	0.8193
3	2	20	−0.0238	0.1766	0.0049	0.2978	0.0843	0.6128
3	2	Varied	−0.0097	0.0999	0.0092	0.1513	0.0592	0.7653
3	2	60	0.0011	0.0930	0.0017	0.1343	0.0501	0.7898
3	2	100	0.0003	0.0736	0.0038	0.1353	0.0340	0.8338
3	10	20	−0.0262	0.1837	0.0030	0.2861	0.1002	0.6092
3	10	Varied	−0.0122	0.1000	−0.0036	0.1486	0.0852	0.7916
3	10	60	−0.0102	0.1007	−0.0063	0.1441	0.0661	0.7768
3	10	100	−0.0054	0.0767	0.0019	0.1155	0.0662	0.8175
20	Varied	20	0.0036	0.0703	0.0226	0.1045	0.0190	0.7850
20	Varied	Varied	0.0006	0.0390	0.0080	0.0576	0.0593	0.8593
20	Varied	60	−0.0004	0.0369	0.0094	0.0555	0.0561	0.8509
20	Varied	100	−0.0003	0.0287	0.0050	0.0402	0.0210	0.9000
20	2	20	−0.0013	0.0685	0.0266	0.1039	0.0633	0.7801
20	2	Varied	0.0021	0.0402	0.0004	0.0568	0.0411	0.8537
20	2	60	0.0035	0.0385	0.0102	0.0520	0.0644	0.8632
20	2	100	−0.0003	0.0286	0.0049	0.0421	0.0483	0.8813
20	10	20	0.0065	0.0663	0.0350	0.1035	0.0220	0.7840
20	10	Varied	−0.0008	0.0385	0.0062	0.0617	0.0522	0.8554
20	10	60	−0.0023	0.0365	0.0072	0.0554	0.0641	0.8707
20	10	100	−0.0027	0.0298	0.0060	0.0476	0.0481	0.8826

behavior of the mean bias and mean absolute deviation as estimated by the Monte Carlos (for example, in the fourth row of Table IV). Given the possible lack of moments of the estimators for these designs, care must be taken in interpreting the columns corresponding to mean bias and mean absolute deviation. On the other hand, poor performance in terms of median bias and median absolute deviation can be interpreted as evidence that the estimators perform poorly.

C.3. *When Is the Single Large Market Limiting Model a Good Approximation?*

The results of Section 3.1 show that, under asymptotics where the number of products increases with firm size and the number of markets fixed, IV estimators do no better than they would with constant markups. To address how well this limiting model approximates Bertrand equilibrium with the Monte Carlo data generating processes used in this section, I simulate from the same data generating process for x, ξ, and MC used in Tables IV and V (which subsume Tables I and III in the main text, but contain results for additional data generating processes as described above), but set markups to $1/\alpha$ for all products and compare estimates based on these data sets to the previously reported estimates computed from data sets with Bertrand prices. Table VI reports the results of applying the same BLP instrument-based estimators to the Monte Carlo data sets with constant markups, while Table VII reports the results for cost instruments.

TABLE VI

MONTE CARLO RESULTS FOR BLP INSTRUMENTS WITH CONSTANT MARKUPS

Markets	Firm Size	Products per Market	Median Bias	Median Abs. Dev. From α_0	Mean Bias	Mean Abs. Dev. From α_0	Rejection Prob. at True α	Power of Test of $\alpha = 0$
1	2	20	−0.3318	0.6416	−0.2259	1.1425	0.1054	0.1486
1	2	60	−0.3589	0.6896	−0.5748	1.5334	0.0842	0.1032
1	2	100	−0.3272	0.6853	−0.6563	1.7149	0.0874	0.1206
1	10	20	1.4864	28.2989	−303.2969	1704.8748	0.3064	0.3064
1	10	60	−0.3112	0.6440	−0.3565	1.4920	0.0521	0.0922
1	10	100	−0.3156	0.6748	−0.5056	1.7151	0.1117	0.1368
3	Varied	20	−0.2828	0.6433	−0.3962	1.3015	0.0130	0.0560
3	Varied	Varied	−0.3300	0.7105	−0.3652	1.5209	0.0110	0.0460
3	Varied	60	−0.3228	0.7043	−0.2699	1.3547	0.0090	0.0590
3	Varied	100	−0.3146	0.6614	−0.3707	1.3190	0.0060	0.0470
3	2	20	−0.3583	0.7749	−0.5273	1.4379	0.0912	0.1082
3	2	Varied	−0.3333	0.6597	−0.3441	1.4748	0.0160	0.0551
3	2	60	−0.3485	0.7714	−0.3107	1.4713	0.0110	0.0591
3	2	100	−0.3118	0.7599	−0.1014	1.7674	0.0340	0.0791
3	10	20	−0.3069	0.7160	−0.3308	1.5446	0.0150	0.0520
3	10	Varied	−0.3049	0.7559	−0.2353	1.4444	0.0090	0.0560
3	10	60	−0.3540	0.7290	−0.3361	1.3365	0.0120	0.0460
3	10	100	−0.3341	0.7353	−0.1354	1.8455	0.0250	0.0581
20	Varied	20	−0.3111	0.7932	−0.6371	2.3960	0.0100	0.0620
20	Varied	Varied	−0.2830	0.7370	−0.1486	1.6991	0.0090	0.0580
20	Varied	60	−0.3471	0.8158	−0.3022	1.9232	0.0080	0.0450
20	Varied	100	−0.3545	0.7563	−0.4122	1.9241	0.0060	0.0530
20	2	20	−0.3432	0.8074	−0.1088	2.1540	0.0150	0.0600
20	2	Varied	−0.3514	0.7758	−0.4193	1.6797	0.0130	0.0570
20	2	60	−0.3504	0.8160	−0.5682	2.2721	0.0060	0.0460
20	2	100	−0.3279	0.8166	−0.2851	1.8619	0.0080	0.0580
20	10	20	−0.3292	0.7525	−0.4875	1.9865	0.0100	0.0430
20	10	Varied	−0.3570	0.8237	−0.4159	1.6799	0.0090	0.0500
20	10	60	−0.3387	0.8265	0.1271	2.3312	0.1533	0.1814
20	10	100	−0.3454	0.7592	−0.2575	2.1509	0.0090	0.0470

TABLE VII

MONTE CARLO RESULTS FOR COST INSTRUMENTS WITH CONSTANT MARKUPS

Markets	Products per Market	Median Bias	Median Abs. Dev. From α_0	Mean Bias	Mean Abs. Dev. From α_0	Rejection Prob. at True α	Power of Test of $\alpha = 0$
1	20	−0.0614	0.3010	0.0221	0.4937	0.1470	0.4673
1	60	−0.0148	0.1538	−0.0067	0.2716	0.0843	0.7329
1	100	−0.0185	0.1233	0.0034	0.1942	0.0604	0.7613
3	20	−0.0100	0.1694	0.0092	0.2583	0.0582	0.6790
3	Varied	−0.0085	0.0934	0.0119	0.1454	0.0654	0.7827
3	60	−0.0099	0.0969	0.0053	0.1378	0.0350	0.7778
3	100	−0.0025	0.0736	0.0143	0.1051	0.0431	0.8317
20	20	0.0039	0.0693	0.0250	0.1027	0.0731	0.7675
20	Varied	−0.0002	0.0371	0.0058	0.0530	0.0581	0.8768
20	60	0.0004	0.0363	0.0142	0.0564	0.0320	0.8639
20	100	−0.0002	0.0282	0.0054	0.0410	0.0230	0.8979

The results show that, while the limiting model gives a pessimistic description of the behavior of BLP instrument-based estimates for some of the cases considered, in other cases it is accurate enough that one would worry about applying the BLP instruments. With a single market, BLP instruments do not appear to perform noticeably better under Bertrand pricing than in the limiting model in any of the Monte Carlo designs. With 3 markets, 10 products per firm, and 100 products, the median bias and median absolute deviation of the estimate of α are only slightly better in the true model than they are with a constant markup, and the size distortion in the two-sided test for α is actually worse. As seen in Table II, the results can be equally bad with 20 markets and 100 products, depending on the ownership structure and coefficient of x in the demand specification.

REFERENCES

BERRY, S., AND A. PAKES (2007): "The Pure Characteristics Demand Model," *International Economic Review*, 48 (4), 1193–1225. [26]

BRESNAHAN, T. F. (1987): "Competition and Collusion in the American Automobile Industry: The 1955 Price War," *The Journal of Industrial Economics*, 35 (4), 457–482. [26]

DEVROYE, L. (1981): "Laws of the Iterated Logarithm for Order Statistics of Uniform Spacings," *The Annals of Probability*, 9 (5), 860–867. [27]

KONOVALOV, A., AND Z. SANDOR (2010): "On Price Equilibrium With Multi-Product Firms," *Economic Theory*, 44 (2), 271–292. [11]

STAIGER, D., AND J. H. STOCK (1997): "Instrumental Variables Regression With Weak Instruments," *Econometrica*, 65 (3), 557–586. [1]

STOCK, J. H., AND J. H. WRIGHT (2000): "GMM With Weak Identification," *Econometrica*, 68 (5), 1055–1096. [2]

VIVES, X. (2001): *Oligopoly Pricing*. MIT Press, Cambridge. [24]

Dept. of Economics and Cowles Foundation, Yale University, 30 Hillhouse Ave., New Haven, CT 06511, U.S.A.; timothy.armstrong@yale.edu.

Guest Co-editor Joris Pinkse handled this manuscript.

Manuscript received February, 2012; final revision received April, 2016.