Supplement to "Large Market Asymptotics for Differentiated Product Demand Estimators with Economic Models of Supply"

Timothy B. Armstrong
Yale University

July 3, 2013

This supplement contains proofs as well as auxiliary results and monte carlos. Section A contains proofs of results in the main text. Section B discusses large market asymptotics in some additional settings. Section C provides primitive conditions for the condition on equilibrium prices in Theorem 1, and relaxes some of the support conditions in that theorem. Section D 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 equivalence results used in the rest of the section. The rest of the section contains proofs of the 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 $\hat{\beta} = \left[\left(\frac{1}{J}\sum_{j=1}^J z_j x_j'\right) W_J \left(\frac{1}{J}\sum_{j=1}^J z_j x_j'\right)\right]^{-1} \left(\frac{1}{J}\sum_{j=1}^J z_j x_j'\right)' W_J \left(\frac{1}{J}\sum_{j=1}^J z_j y_j'\right)$ 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 the following, 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)* $\sqrt{J}\left(\frac{1}{J}\sum_{j=1}^{J} z_j x_j' - M_{zx}\right) \overset{d}{\to} Z_{zx}$ *for some* $k \times d$ *matrix* M_{zx}, *and a* $k \times d$ *random matrix* Z_{zx}. *(ii)* $\frac{1}{\sqrt{J}}\sum_{j=1}^{J} z_j \xi_j \overset{d}{\to} Z_{z\xi}$ *for a multivariate normal random vector* $Z_{z\xi}$ *with nonsingular variance. (iii)* $W_J \overset{p}{\to} W$ *for some positive definite weighting matrix* W.

Assumption 2. $\sqrt{J}\max_j \|x_j^* - x_j\| \overset{p}{\to} 0$ *and* $\frac{1}{J}\sum_{j=1}^{J}\|z_j\| = \mathcal{O}_P(1)$.

Theorem 4. *Under Assumption 1, we have the following.*

(i) Let $d_2 = \operatorname{rank} M_{zx}$ *and* $d_1 = d - d_2$. *Let* H *be an invertible* $d \times d$ *matrix such that the first* d_1 *columns of* $M_{zx}H$ *are zero and split* H *into its first* d_1 *and last* d_2 *columns as* (H_1, H_2). *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*

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

where $Q_{W,1}$ *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* $Ezx'H_2$.

(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, $\|\hat{\beta} - \hat{\beta}^*\| \overset{p}{\to} 0$.

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_2 diagonal entries equal to 1 and the last d_1 equal to $\sqrt{J}$,

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

so that

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

By the continuous mapping theorem, this converges to

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

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

For part (ii), the note that, under Assumptions 1 and 2, Assumption 1 will also hold with x_j^* and y_j^*. In fact, we will have $\left(\sqrt{J}\left(\frac{1}{J}\sum_{j=1}^J z_j x_j^{*\prime} - M_{zx}\right), \sqrt{J}\left(\frac{1}{J}\sum_{j=1}^J z_j x_j' - M_{zx}\right)\right) \stackrel{d}{\to}$ (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. □

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 4 (with M_{zx} full rank).

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

Assumption 4. $\max_j \|x_j^* - x_j\| \stackrel{p}{\to} 0$ *and* $\frac{1}{J}\sum_{j=1}^J \|z_j\| = \mathcal{O}_P(1)$.

Theorem 5. *Under Assumptions 3 and 4,* $\sqrt{J}(\hat{\beta} - \beta)$ *and* $\sqrt{J}(\hat{\beta}^* - \beta)$ *are consistent and asymptotically normal, with the same asymptotic distribution.*

Proof. Under these assumptions, Assumption 3 holds for both the original and starred quantities. The result then follows from standard arguments. □

The following corollary restates some immediate implications of the above theorems in a more concise way, for use in the main text.

Corollary 1. *If Assumptions 1 and 2 hold, $\hat{\beta}$ is consistent iff. $\hat{\beta}^*$ is consistent (iff M_{zx} is full rank) and, in particular, $\|\hat{\beta} - \hat{\beta}^*\| \xrightarrow{p} 0$. In addition, if $\hat{\beta}^*$ is consistent and asymptotically normal, then $\hat{\beta}$ is consistent and asymptotically normal with the same asymptotic distribution.*

A.2 Proof of Theorem 1

By the central limit theorem, Assumption 1 holds with Z_{zx} and $Z_{z\xi}$ normal random variables. Thus, the result follows from Corollary 1 by verifying Assumption 2 for (x_j, p_j^*).

The markup can be written as

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

Thus, it suffices to show that $\left(\int \tilde{s}_j^2(\delta,\zeta)\, dP_\zeta(\zeta)\right) / \left(\int \tilde{s}_j(\delta,\zeta)\, dP_\zeta(\zeta)\right)$ converges to zero at a faster than $\sqrt{J}$ rate. Fix a sequence $k_J \to \infty$. Let $\overline{s}_J$ and $\underline{s}_J$ be the supremum and infimum respectively of $\tilde{s}_j(\delta,\zeta)$ with $\|\zeta\| \le k_J$, $j \le J$, and the elements of δ ranging over the given bounded set. Then

$$\begin{aligned}
&\max_{j\le J} \frac{\int \tilde{s}_j^2(\delta,\zeta)\, dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\, dP_\zeta(\zeta)} \le \max_{j\le J} \frac{\int \tilde{s}_j^2(\delta,\zeta) I(\|\zeta\| \le k_J)\, dP_\zeta(\zeta) + P_\zeta(\|\zeta\| > k_J)}{\int \tilde{s}_j(\delta,\zeta) I(\|\zeta \le k_J\|)\, dP_\zeta(\zeta)} \\
&\le \frac{\int \overline{s}_J^2 I(\|\zeta\| \le k_J)\, dP_\zeta(\zeta) + P_\zeta(\|\zeta\| > k_J)}{\int \underline{s}_J I(\|\zeta \le k_J\|)\, dP_\zeta(\zeta)} \le \frac{\overline{s}_J^2 + P_\zeta(\|\zeta\| > k_J)}{\underline{s}_J(1 - P_\zeta(\|\zeta\| > k_J))}.
\end{aligned}$$

If we can choose k_J so that $\sqrt{J} P_\zeta(\|\zeta\| > k_J)/\underline{s}_J$ and $\sqrt{J}\overline{s}_J^2/\underline{s}_J$ both go to zero, we will have the desired result. Since product characteristics are bounded, there exists some B such that, for all j,

$$\left|\sum_k x_{jk}\zeta_k\right| \le B\|\zeta\|.$$

Letting M be a bound for δ_j, this gives the following bounds for $\overline{s}_J$ and $\underline{s}_J$:

$$\begin{aligned}
\overline{s}_J &\le \frac{\exp(M + Bk_J)}{\sum_\ell \exp(-M - Bk_J)} = \frac{\exp(2M + 2Bk_J)}{J} \\
\underline{s}_J &\ge \frac{\exp(-M - Bk_J)}{\sum_\ell \exp(M + Bk_J)} \ge \frac{\exp(-2M - 2Bk_J)}{J}.
\end{aligned}$$

This gives $\bar{s}_J^2/\underline{s}_J \leq \exp(6M + 6Bk_J)/J$. If the distribution of ζ is joint normal, then, for some constants K_1 and K_2,

$$P_\zeta(\|\zeta\| > k_J) \leq K_1 \exp(-K_2 k_J^2).$$

If this holds, then

$$\frac{P_\zeta(\|\zeta\| > k_J)}{\underline{s}_J} \leq \frac{K_1 \exp(-K_2 k_J^2) J}{\exp(-2M - Bk_J)} = K_1 \exp(-K_2 k_J^2 + Bk_J + 2M) J.$$

For $k_J = (\log J)^{2/3}$, we will have $\sqrt{J}\bar{s}_J^2/\underline{s}_j \to 0$ and $\sqrt{J}P_\zeta(\|\zeta\| > k_J)/\underline{s}_J \to 0$ as desired.

A.3 Proof of Theorem 2

If $\frac{1}{J}\sum_{j=1}^J Ez_j(x_j', p_j^*)$ converges in probability to a positive definite matrix, then, by a law of large numbers for inid variables, $\frac{1}{J}\sum_{j=1}^J z_j(x_j', p_j^*)$ will converge in probability to the same matrix. The result then follows by Theorem 5 as long as $\max_j \|p_j - p_j^*\| \stackrel{p}{\to} 0$.

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 continuously differentiable at $(\pi_1\mu_{r,1}, \ldots, \pi_F\mu_{r,F})$ (the latter claim can be seen using the implicit function theorem). The difference between p_j and p_j^* can then be written as, for f the firm producing product j, $b_f(\pi_1\mu_{r,1}, \ldots, \pi_F\mu_{r,F}) - b_f(\hat{\pi}_1\bar{r}_1, \ldots, \hat{\pi}_F\bar{r}_F)$, which converges in probability to zero by the law of large numbers. Since $\max_j \|p_j - p_j^*\| = \max_f b_f(\pi_1\mu_{r,1}, \ldots, \pi_F\mu_{r,F}) - b_f(\hat{\pi}_1\bar{r}_1, \ldots, \hat{\pi}_F\bar{r}_F)$ and the number of firms does not increase with J, the result follows.

A.4 Proof of Theorem 3

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 $m_2 = \frac{1}{N}\sum_{i=1}^N (J_i/\bar{J})^2$, $m_3 = \frac{1}{N}\sum_{i=1}^N (J_i/\bar{J})^3$, $m_{2,\infty} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N (J_i/\bar{J})^2$ 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 $\mu_r = E(r_{i,j})$, $\mu_{xr} = E(x_{i,j}r_{i,j})$ and $\mu_{wr} = E(w_{i,j}r_{i,j})$.

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

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

and let $u_{i,j} = \xi_{i,j}(x'_{i,j}, \mu'_h)'$. Let Σ_{Wu} be the variance matrix of $(vec(W_{i,j})', u'_{i,j})'$ 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 $3d_z$, etc. (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

$$M_1 = \begin{pmatrix} I_{d_x} & 0 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} E(x_{i,j}x'_{i,j}) & E(x_{i,j}(MC_{i,j} + \frac{1}{\alpha})) \\ 0 & 1 \end{pmatrix}, \quad K_{m_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & I_{d_x - 1} & 0 \\ m_2\mu_h & 0 & 1_{d_h \times 1} \end{pmatrix}$$

where $1_{d_h \times 1}$ is a $d_h \times 1$ vector of ones, and let $K_{m_{2,\infty}}$ be defined in the same way, but with $m_{2,\infty}$ replacing m_2. Let

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

for any positive real number q. Let

$$\tilde{M}_1 = \begin{pmatrix} 1 & E(w_{i,j})' & E(MC_{i,j}) + \frac{1}{\alpha} \\ E(w_{i,j}) & E(w_{i,j}w'_{i,j}) & E(w_{i,j}MC_{i,j}) + \frac{1}{\alpha}E(w_{i,j}) \\ \mu_h m_2 & \mu_h m_2 E(w_{i,j})' & \mu_h m_2 \left[E(MC_{i,j}) + \frac{1}{\alpha}\right] \end{pmatrix}$$

Note that $A_q^{-1} = A_{1/q}$ and, with the above notation, $\tilde{M}_1 = K_{m_2} M_1 H$.

We first prove the following lemma.

Lemma 1. *Under the conditions of Theorem 3,*

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

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

Proof. First, note that

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

Now, substituting the bound $0 \leq p_{i,j} - MC_{i,j} - 1/\alpha \leq C/(J\alpha)$ on the right hand side, which holds for C large enough, we see that the above display is bounded from above by

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

This, and a similar lower bound give the result with the sum in the denominator replaced by the same sum over $k\neq j$. The result then follows since $r_{i,j}/\sum_{k\neq j}r_{i,k}=(r_{i,j}/\sum_{k=1}^{J_i}r_{i,k})(1+r_{i,j}/\sum_{k\neq j}r_{i,k})$, which is equal to $r_{i,j}/\sum_{k=1}^{J_i}r_{i,k}$ plus a term that is bounded by a constant times $1/J^2$. □

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

Lemma 2. *Under the conditions of Theorem 3,*

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

Proof. We have

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

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

Lemma 3. *Under the conditions of Theorem 3,*

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

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 1. The last term is bounded by a constant times $1/J_i$. For the first term, we have

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

The last two terms converge in probability to zero (this can be shown by bounding their second moments). □

Lemma 4. *Under the conditions of Theorem 3, for any sequence of iid variables $v_{i,j}$ with mean μ_v and a finite fourth moment,*

$$\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left[\frac{1}{\bar{J}}\sum_{k=1}^{J_i} h(x_{i,k})\right] v_{i,j}$$
$$= m_2\mu_h\mu_v + \frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}\left[\mu_h(v_{i,j} - \mu_v) + \mu_v(h(x_{i,j}) - \mu_h)\right] + o_P\left(1/\sqrt{N\bar{J}}\right).$$

Proof. We have

$$
\begin{aligned}
&\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\left[\frac{1}{\bar{J}}\sum_{k=1}^{J_i}h(x_{i,k})\right]v_{i,j}\\
&=\mu_h\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}v_{i,j}+\frac{1}{N\bar{J}}\sum_{i=1}^{N}\left\{\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\right\}\left[\frac{1}{\bar{J}}\sum_{j=1}^{J_i}v_{i,j}\right]\\
&=\mu_h\mu_v\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}+\mu_h\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}[v_{i,j}-\mu_v]\\
&+\mu_v\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\frac{J_i}{\bar{J}}+\frac{1}{N\bar{J}}\sum_{i=1}^{N}\left\{\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\right\}\left\{\frac{1}{\bar{J}}\sum_{j=1}^{J_i}[v_{i,j}-\mu_v]\right\}\\
&=\mu_h\mu_v\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}+\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}\left\{\mu_h[v_{i,j}-\mu_v]+\mu_v[h(x_{i,j})-\mu_h]\right\}\\
&+\frac{1}{N\bar{J}}\sum_{i=1}^{N}\left\{\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\right\}\left\{\frac{1}{\bar{J}}\sum_{j=1}^{J_i}[v_{i,j}-\mu_v]\right\}\\
&\equiv I+II+III
\end{aligned}
$$

where $I \equiv \mu_h\mu_v m_2$, $II \equiv \frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\frac{J_i}{\bar{J}}\left\{\mu_h[v_{i,j}-\mu_v]+\mu_v[h(x_{i,j})-\mu_h]\right\}$ and $III \equiv \frac{1}{N\bar{J}}\sum_{i=1}^{N}\left\{\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\right\}\left\{\frac{1}{\bar{J}}\sum_{j=1}^{J_i}[v_{i,j}-\mu_v]\right\}$. Note that $III=\mathcal{O}_P(1/(\bar{J}\sqrt{N}))$ since

$$
var(III)=\frac{1}{N\bar{J}^2}\frac{1}{N}\sum_{i=1}^{N}var\left(\left\{\frac{1}{\sqrt{\bar{J}}}\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\right\}\left\{\frac{1}{\sqrt{\bar{J}}}\sum_{j=1}^{J_i}[v_{i,j}-\mu_v]\right\}\right)
$$

and the variance on the right hand side can be seen to be bounded uniformly in J_i by an application of Cauchy-Schwarz and boundedness of the fourth moments of $\left\{\frac{1}{\sqrt{\bar{J}}}\sum_{j=1}^{J_i}[h(x_{i,j})-\mu_h]\right\}$ and $\left\{\frac{1}{\sqrt{\bar{J}}}\sum_{j=1}^{J_i}[v_{i,j}-\mu_v]\right\}$. □

Lemma 5. *Under the conditions of Theorem 3,*

$$
\frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}\begin{pmatrix}x_{i,j}\\ \frac{1}{\bar{J}}\sum_{k=1}^{J}h(x_{i,k})\end{pmatrix}\begin{pmatrix}x'_{i,j} & MC_{i,j}+\frac{1}{\alpha}\end{pmatrix}=\tilde{M}_1+V_{JN}/\sqrt{N\bar{J}}
$$

where $\tilde{M}_1$ is defined at the beginning of this section and $\left((vec(V_{JN})', \left(\frac{1}{\sqrt{N\bar{J}}}\sum_{i=1}^{N}\sum_{j=1}^{J_i}z_{i,j}\xi_{i,j}\right)'\right)'$ 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 4 that $\left((vec(V_{JN})', \left(\frac{1}{\sqrt{N\bar{J}}}\sum_{i=1}^{N}\sum_{j=1}^{J_i} z_{i,j}\xi_{i,j}\right)'\right)'$ is, up to $o_P(1)$, equal to

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

where, for any scalar r, B_r is defined to be the $d_z \times d_z$ diagonal matrix with ones in the first d_x diagonal entries and r in the remaining diagonal entries. By a central limit theorem for triangular arrays of inid 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.

□

Putting the above lemmas together and using the fact that $\tilde{M}_1 = K_{m_2}M_1H$, we have

$$\hat{M}_{zx} \equiv \frac{1}{N\bar{J}}\sum_{i=1}^{N}\sum_{j=1}^{J_i} z_j(x_j', p_j) = K_{m_2}M_1H + V_{JN}/\sqrt{\bar{J}N} + M_2(A_{1/\bar{J}} + o_P(1/\bar{J}))$$

where V_{JN} is given in Lemma 5. Since the last column of M_1 is all zeros, $M_1A_{\bar{J}} = M_1$. Also, since the first d_x columns of $M_2A_{\bar{J}}$ are zero, $M_2A_{\bar{J}}H = M_2A_{\bar{J}}$, so that $M_2A_{\bar{J}}H^{-1} = M_2A_{\bar{J}}$ as well. Thus,

$$Q_N \equiv \hat{M}_{zx}H^{-1}A_{1/\bar{J}}^{-1} = K_{m_2}M_1 + V_{JN}H^{-1}A_{\bar{J}}/\sqrt{\bar{J}N} + M_2 + o_P(1).$$

Note that $K_{m_2}M_1 + M_2$ is full rank iff $m_2 \neq 1$, so that this matrix is full rank under the conditions of the theorem. Let $\hat{Z}_{z\xi} = \frac{1}{\sqrt{\bar{J}N}}\sum_{i=1}^{N}\sum_{j=1}^{J_i} z_{i,j}\xi_{i,j}$ (which converges jointly with V_{JN} to a vector of normal variables).

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

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

where $Q = K_{m_2,\infty}M_1 + M_2$ and $Z_{z\xi}$ is a normal vector defined below. In the case where $N/\bar{J} \to c$ for a finite constant c,

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

This converges in distribution to $K_{m_2,\infty}M_1 + Z_{zx}H^{-1}diag(0,\ldots,0,1) + M_2A_{\sqrt{c}} \equiv \tilde{Q}_\infty$ jointly with $\hat{M}_{z\xi}$, where $(vec(Z_{zx})', Z'_{z\xi})'$ is normal with mean zero and variance matrix $\widetilde{\Sigma}_{Wu}$ (so that $\hat{M}_{z\xi}$ and V_{JN} converge in distribution jointly to $Z_{z\xi}$ and Z_{zx} by Lemma 5). Thus,

$$\begin{aligned} &diag(\sqrt{\bar{J}N},\ldots,\sqrt{\bar{J}N},1)H[(\hat{\beta}',\hat{\alpha})' - (\beta',\alpha)'] = \sqrt{\bar{J}N}A_{1/\sqrt{N\bar{J}}}H[(\hat{\beta}',\hat{\alpha})' - (\beta',\alpha)'] \\ &= \left(\tilde{Q}'_N W\tilde{Q}_N\right)^{-1}\tilde{Q}'_N W\hat{Z}_{z\xi} \overset{d}{\to} \left(\tilde{Q}'_\infty W\tilde{Q}_\infty\right)^{-1}\tilde{Q}'_\infty W Z_{z\xi}. \end{aligned}$$

For $c = 0$, $\tilde{Q}_\infty = K_{m_2,\infty}M_1 + Z_{zx}H^{-1}diag(0,\ldots,0,1)$, 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 multiproduct 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 $g \in \{1,\ldots,G\}$ is denoted by $\mathcal{J}_g \subseteq \{1,\ldots,J\}$. 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

$$\log s_j - \log s_0 = x_j'\beta - \alpha p_j + \sigma \log \bar{s}_{j/g} + \xi_j \tag{8}$$

(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

$$p_j - MC_j = \frac{1-\sigma}{\alpha} / (1 - \sigma\bar{s}_{j/g} - (1-\sigma)s_j). \tag{9}$$

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 (9). 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

$$\begin{aligned} p_j^* - MC_j &= \frac{1-\sigma}{\alpha} / (1 - \sigma\bar{s}_{j/g}(x, p^*, \xi)) \\ &= \frac{1-\sigma}{\alpha} \frac{\sum_{k\in\mathcal{J}_g} \exp((x_k'\beta - p_k^*\alpha + \xi_k)/(1-\sigma))}{\left[\sum_{k\in\mathcal{J}_g} \exp((x_k'\beta - p_k^*\alpha + \xi_k)/(1-\sigma))\right] - \sigma \exp((x_j'\beta - p_j^*\alpha + \xi_j)/(1-\sigma))} \end{aligned} \tag{10}$$

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 (9), 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 6. *In the nested logit model single product firms and many nests, suppose that* (x_j, ξ_j, MC_j) *is bounded and iid across* j. *Let* $z_j = (x_j, h(\{x_k\}_{j\in\mathcal{J}_{g-L}}, \ldots, \{x_k\}_{j\in\mathcal{J}_{g+M}})$ *for* $j \in \mathcal{J}_g$ *for some function* h *with finite variance. Let* p_j^* *and* $\bar{s}_{j/g}^*$ *be defined in (10). 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 (9), 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 6 is given below.

proof of Theorem 6. 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 $\sqrt{J}$ rate.

Define $f : \mathbb{R}^{(4+d)|\mathcal{J}_g|+2d+2} \to \mathbb{R}^{|\mathcal{J}_g|}$ by

$$\begin{aligned} &f_j(p, x, \xi, \eta, \theta, r) \\ &= p_j - MC_j(x, \eta, \theta) \\ &- \frac{1-\sigma}{\alpha} \frac{\sum_k \exp((x_k'\beta - p_k\alpha + \xi_k)/(1-\sigma))}{[\sum_k \exp((x_k'\beta - p_k\alpha + \xi_k)/(1-\sigma))] - \sigma \exp((x_j'\beta - p_j\alpha + \xi_j)/(1-\sigma))} + r_j. \end{aligned}$$

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

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

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 $f(p, x, \xi, \eta, r) = 0$ 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 and $MC(x, \theta, \eta)$ bounded away from zero, the equation $f(p, x, \theta, \eta, r) = 0$ 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)):

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

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

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

and, for $k \neq j$,

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

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 $f(p, x, \xi, \eta, \theta, r) = 0$ for the elements of (x, ξ, η) in the given bounded set, θ in the given neighborhood of θ_0, and r close to zero, this defines p as a function $\phi(x, \xi, \eta, \theta, r)$ of the remaining variables. By the implicit function theorem, the derivative matrix of ϕ is given by

$$D\phi(x, \xi, \eta, \theta, r) = (D_p f(\phi(x, \xi, \eta, \theta, r), x, \xi, \eta, \theta, r))^{-1} D_{x,\xi,\eta,\theta,r} f(\phi(x, \xi, \eta, \theta, r), x, \xi, \eta, \theta, r)$$

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 $p = \phi(x, \xi, \eta, \theta, \tilde{r})$ and $p^* = \phi(x, \xi, \eta, \theta, 0)$, by the mean value theorem, for every index j, there is a $\overline{r}$ between 0 and $\tilde{r}$ such the difference between p_j and p_j^* is given by the jth row of

$$(D_p f(\phi(x, \xi, \eta, \theta, \overline{r}), x, \xi, \eta, \theta, \overline{r}))^{-1} D_r f(\phi(x, \xi, \eta, \theta, \overline{r}), x, \xi, \eta, \theta, \overline{r})\tilde{r}.$$

Since the elements of $(D_p f(\phi(x, \xi, \eta, \theta, r), x, \xi, \eta, \theta, r))^{-1} D_r f(\phi(x, \xi, \eta, \theta, r), x, \xi, \eta, \theta, r)$ are continuous functions of x, ξ, η, θ, and r, the function that maps t to the maximum of the absolute values of the elements of $(D_p f(\phi(x, \xi, \eta, \theta, r), x, \xi, \eta, \theta, r))^{-1} D_r f(\phi(x, \xi, \eta, \theta, r), x, \xi, \eta, \theta, r)t$ takes a maximum M as x, ξ, η, θ, 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

$$\sqrt{J} \max_{j \leq J, \|\theta - \theta_0\| < \varepsilon} |p_j^* - p_j| \leq \sqrt{J} \max_{j \leq J} M \|\tilde{r}_j\| \to 0.$$

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 $(x, \xi, \eta, \theta, p)$ that can be taken under the assumptions of the theorem. Thus, by the mean value theorem, for some finite B, $\sqrt{J} \max_{j \leq J, \|\theta - \theta_0\| < \varepsilon} |\bar{s}_{j/g}^* - \bar{s}_{j/g}| \leq \sqrt{J} B \max_{j \leq J, \|\theta - \theta_0\| < \varepsilon} |p_j^* - p_j| \to 0$. □

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). 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 $\delta_1 < \ldots < \delta_J$. If all products have positive market share, this will imply that prices satisfy $p_1 < \ldots < p_J$ as well. Consumer i will prefer good j to $j-1$ if

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

Combining this with the expression for $j+1$, consumer i will prefer j to its neighbors if $\Delta_j > \zeta_{ip} > \Delta_{j+1}$. In order for all products to have positive market share, this must hold for some ζ_{ip} for all j, so we must have $\Delta_1 > \ldots > \Delta_J$. 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 cdf of ζ_{ip}, market shares will be given by

$$s_j = F(\Delta_j) - F(\Delta_{j+1}). \tag{11}$$

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

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

If F is known, this equation can be estimated using OLS. If F is allowed to depend on an unknown parameter, more instruments will be needed, so it will be useful to study the identifying power of moment conditions based on characteristics of other products.

Differentiating the formula for shares with respect to p_j gives, letting f be the pdf 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

$$p_j - MC_j = \frac{F(\Delta_j) - F(\Delta_{j+1})}{f(\Delta_j)\frac{\Delta_j}{p_j - p_{j-1}} + f(\Delta_{j+1})\frac{\Delta_{j+1}}{p_{j+1} - p_j}}. \tag{12}$$

Suppose that, for some $\underline{\zeta} > 0$, $\underline{\zeta} \leq \zeta_{ip}$ 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 Equation 12 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 4 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 $\sqrt{J}\max_{j\leq J}\delta_j - \delta_{j-1} \to 0$ (e.g., this holds with probability one 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

$$p_j - MC_j = \frac{F(\Delta_j) - F(\Delta_{j+1})}{f(\Delta_j)\frac{\Delta_j}{p_j - p_{j-1}} + f(\Delta_{j+1})\frac{\Delta_{j+1}}{p_{j+1} - p_j}} \leq \frac{\overline{f}}{\underline{f}} \frac{\Delta_j - \Delta_{j+1}}{\frac{\Delta_j}{p_j - p_{j-1}} + \frac{\Delta_{j+1}}{p_{j+1} - p_j}} \leq \frac{\overline{f}}{\underline{f}}(p_j - p_{j-1})$$

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

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

Thus,

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

B.3 Multi Product Firms

This section considers the case with many small multiproduct firms. If the number of products sold by each firm is fixed and the number of firms grows large, the results are similar 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

$$(p_j - MC_j)(1 - s_j(x,p,\xi)) - \sum_{k \in \mathcal{F}_f, k \neq j} (p_k - MC_k) s_k(x,p,\xi) - \frac{1}{\alpha} = 0. \tag{13}$$

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

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

Rearranging gives

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

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

$$0 = \frac{1-\sigma}{\alpha} - (p_j - MC_j)(1 - \sigma \bar{s}_{j/g}) + \sum_{k \in \mathcal{F}_f \cap \mathcal{J}_g, k \neq j} (p_k - MC_k)\sigma \bar{s}_{k/g} + \tilde{r}_J. \tag{14}$$

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 the one used for Theorem 6 will show that prices in the nested logit game converge uniformly at a faster than $\sqrt{J}$ 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

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

This can be rearranged to give

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

Under the assumptions of Theorem 1, the left hand side converges to $(p_j - MC_j)$ at faster than a $\sqrt{J}$ 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 $\frac{\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}$, since, using the notation of the proof of Theorem 1, $\int \tilde{s}_j(\delta,\zeta)\tilde{s}_k(\delta,\zeta)\,dP_\zeta(\zeta) \leq \overline{s}_J^2 + P_\zeta(\|\zeta\| > k_J)$, giving the same bound on the numerator.

C Support Conditions

This section provides primitive sufficient conditions for the bounded price assumption in Theorem 1, and relaxes the assumption of bounded product characteristics. Theorem 7 shows that the high level assumption on prices in Theorem 1 will hold as long as the support of the random coefficients $\|\zeta\|$ is bounded. Theorem 8 relaxes the finite support condition for product characteristics to an exponential tail bound, also using an assumption of bounded random coefficients. It seems likely that some of these conditions could be relaxed further using more involved arguments, although this is left for future research.

Theorem 7. *In the random coefficients logit model, suppose that the support of the distribution of the random coefficients ζ is bounded and that the sequence of marginal costs and* *product characteristics is bounded. Then, for any Nash-Bertrand equilibrium of the pricing game with single product firms, prices are bounded by a constant that does not depend on J.*

Proof. For B a bound on $x_k'(\beta+\zeta)+\xi_k$ over all j and the support of ζ,

$$\frac{\exp(-2B-\alpha p_j)}{\sum_k \exp(-\alpha p_k)} = \frac{\exp(-B-\alpha p_j)}{\sum_k \exp(B-\alpha p_k)} \leq \tilde{s}_j(\delta,\zeta) \leq \frac{\exp(B-\alpha p_j)}{\sum_k \exp(-B-\alpha p_k)} = \frac{\exp(2B-\alpha p_j)}{\sum_k \exp(-\alpha p_k)}$$

for all ζ its support. In particular, $\sup_{\zeta\in\text{supp}(\zeta)} s_j(\delta,\zeta)/\inf_{\zeta\in\text{supp}(\zeta)} s_j(\delta,\zeta) \leq \exp(4B)$. Plugging these bounds into the first order conditions for product j gives

$$p_j - MC_j - \frac{1}{\alpha} = \frac{1}{\alpha}\frac{\int \tilde{s}_j^2(\delta,\zeta)\,dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)(1-\tilde{s}_j(\delta,\zeta))\,dP_\zeta(\zeta)} \leq \exp(4B)\frac{1}{\alpha}\frac{\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}{1-\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)}.$$

The right hand side is increasing in $\int \tilde{s}_j(\delta,\zeta)\,dP_\zeta(\zeta)$, which is bounded by $\exp(B-\alpha p_j)/[1+\exp(B-\alpha p_j)]$ (since one of the other products is the outside good with utility 0), giving a bound of

$$\exp(4B)\frac{1}{\alpha}\frac{\frac{\exp(B-\alpha p_j)}{1+\exp(B-\alpha p_j)}}{1-\frac{\exp(B-\alpha p_j)}{1+\exp(B-\alpha p_j)}} = \exp(4B)\frac{1}{\alpha}\exp(B-\alpha p_j) = \frac{1}{\alpha}\exp(5B-\alpha p_j)$$

for the previous display. Taking logs and rearranging,

$$\alpha p_j \leq -\log\alpha + 5B - \log\left(p_j - MC_j - \frac{1}{\alpha}\right)$$

If $p_j - MC_j - \frac{1}{\alpha} \geq 1$, then the right hand side is less than $-\log\alpha + 5B$, so $p_j \leq (-\log\alpha + 5B) \vee (MC_j + 1/\alpha + 1)$. □

Theorem 8. *Under the setup of Theorem 1, suppose that $\|\zeta\|$ has finite support. Then the condition that the support of prices, product characteristics and marginal costs is bounded can be replaced by the condition that, for some K and $\varepsilon > 0$, $P(\|x_j,\xi_j,MC_j\| > t) \leq K\exp(-t^{1+\varepsilon})$, and the conclusion of the theorem will still hold.*

Proof. If ζ has bounded support and x_k and ξ_k are bounded by some B_J, then, for some constant K_1, $\|x_k'(\beta+\zeta)+\xi_k\| \leq K_1 B_J$ on the support of ζ. In addition, by the bound in Theorem 7, if MC_k is bounded by B_J as well, we will have, for some constant K_2, $p_k \leq K_2 B_J$ for all k.

Thus, on the event that $\|(x_k,\xi_k,MC_k)\| \leq B_J$ for all $k \leq J$, we will have, for a constant K that depends only on K_1, K_2 and α, $\|x_k'(\beta+\zeta)-\alpha p_k+\xi_k\| \leq KB_J$, so, on the support of ζ,

$$\frac{\exp(-2KB_J)}{J} = \frac{\exp(-KB_J)}{\sum_k \exp(KB_J)} \leq \tilde{s}_j(\delta,\zeta) \leq \frac{\exp(KB_J)}{\sum_k \exp(-KB_J)} = \frac{\exp(2KB_J)}{J}.$$

It follows that

$$\frac{\int \tilde{s}_j^2(\delta,\zeta)\, dP_\zeta(\zeta)}{\int \tilde{s}_j(\delta,\zeta)\, dP_\zeta(\zeta)} \leq \frac{\left[\frac{\exp(2KB_J)}{J}\right]^2}{\frac{\exp(-2KB_J)}{J}} = \frac{1}{J}\exp(6KB_J)$$

on this event. Since the markup goes to zero at the same rate as the quantity in the above display, the result will hold as long as the bound on the right hand side decreases at a faster than $\sqrt{J}$ rate, and the event $\|(x_k, \xi_k, MC_k)\| \leq B_J$ for all $k \leq J$ holds with probability approaching one.

The right hand side of the above display decreases more quickly than $1/\sqrt{J}$ for $B_J = (\log J)/(12K+1)$, so the result will follow as long as $P\left(\|(x_j, \xi_j, MC_j)\| > B_J \text{ some } j \leq J\right) \to 0$ for this choice of B_J. Indeed,

$$P\left(\|(x_j, \xi_j, MC_j)\| > B_J \text{ some } j \leq J\right) \leq JP\left(\|(x_1, \xi_1, MC_1)\| > B_J\right) \leq JK \exp\left(-B_J^{1+\varepsilon}\right)$$
$$= JK \exp\left\{-[(\log J)/(12K+1)]^{1+\varepsilon}\right\} \to 0,$$

giving the result. □

D Monte Carlo

This section reports details for the monte carlos, as well as results for designs not reported in the main text. The data generating process for the monte carlo data sets is as follows. x_j contains a constant and a uniform $(0,1)$ random variable. I generate the cost shifter, z_j, as another uniform random variable independent of x. Marginal cost is given by $MC_j = (x_j', z_j')\gamma + \eta_j$ for a η_j defined as follows. To generate η and ξ, I generate three independent uniform $(0,1)$ random variables u_{1j}, u_{2j}, and u_{3j}, and set $\xi_j = .9 \cdot u_{1j} + .1 \cdot u_{3j} - 1$ and $\eta_j = .9 \cdot u_{1j} + .1 \cdot u_{2j} - 1$ so that η and ξ are correlated and bounded. x_j, ξ_j, and η_j are independent across products j. Utility is given by the random coefficients model of section 3.1, with the random coefficient on the covariate given by a $N(0, \sigma^2)$ random variable, where σ^2 is set to 9 and is estimated in the monte carlos. The parameters are given by $\alpha = 1$, $\beta = (3,6)'$, and $\gamma = (2,1,1)'$, where the last element of γ is the coefficient of the excluded cost instrument. Note that with these parameter values, the variance of the observed portion of utility $x'\beta$ across products and the idiosyncratic component ε_{ij} across consumers is of the same order of magnitude (the variance of the extreme value error ε_{ij} is $\pi^2/6$, while the nonconstant element of $x'\beta$ has a variance of 3). 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.).

In addition to the results reported in the main text, which use 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. Finally, I report the results of the first stage F test and the test statistic suggested in section 4.2 (labeled "*BLP F*") applied to the linear part of the model. Note that neither of these tests gives a full test for lack of identification, since they only apply to the linear part of the model.

Regarding the results for the first stage statistics, note that the usual first stage F statistic and the "*BLP F*" statistic perform similarly. This is likely because the markups used for the "*BLP F*" statistic are correctly specified (recall that the "*BLP F*" is designed not to reject too often in cases where the supply side model is misspecified in ways that constrain BLP instruments to perform poorly). From these monte carlos, it appears that the F above 10 rule or a similar rule based on the "*BLP F*" statistic does a decent job of predicting whether the BLP instruments perform well, despite being based on a setting where nonlinear parameters do not have to be estimated (although this finding should be interpreted with caution given the limited scope of the monte carlos).

markets	firm size	products per market	median bias	median absolute deviation from α_0	rejection prob. at true α	power of test of $\alpha = 0$
1	2	20	-0.3385	0.6081	0.1439	0.2052
1	2	60	-0.3613	0.6660	0.0631	0.0731
1	2	100	-0.3491	0.6825	0.1266	0.1628
1	10	20	-0.2147	1.9530	0.2729	0.2729
1	10	60	-0.3698	0.6691	0.0783	0.1004
1	10	100	-0.3648	0.7177	0.1211	0.1381
3	varied	20	-0.0229	0.1665	0.0450	0.7390
3	varied	varied	-0.0890	0.2786	0.0520	0.4700
3	varied	60	-0.0804	0.3922	0.1002	0.2956
3	varied	100	-0.1586	0.4504	0.0160	0.1590
3	2	20	-0.2893	0.6742	0.0280	0.0750
3	2	varied	-0.3313	0.6753	0.0250	0.0600
3	2	60	-0.3697	0.7407	0.0090	0.0530
3	2	100	-0.3154	0.7171	0.0140	0.0600
3	10	20	-0.1053	0.3358	0.0390	0.3980
3	10	varied	-0.0494	0.2966	0.1523	0.4649
3	10	60	-0.2186	0.5827	0.0200	0.1040
3	10	100	-0.2525	0.6383	0.1351	0.1762
20	varied	20	-0.0044	0.0504	0.0510	1.0000
20	varied	varied	-0.0211	0.1537	0.0480	0.9170
20	varied	60	-0.0061	0.1158	0.0400	0.9990
20	varied	100	-0.0190	0.1659	0.0410	0.9450
20	2	20	-0.0393	0.3504	0.1552	0.4535
20	2	varied	-0.1578	0.4697	0.0851	0.2543
20	2	60	-0.1689	0.6458	0.0090	0.1080
20	2	100	-0.2191	0.6897	0.1061	0.1632
20	10	20	0.0039	0.1140	0.0390	0.9880
20	10	varied	-0.0014	0.1001	0.0400	0.9960
20	10	60	0.0130	0.2345	0.0230	0.7710
20	10	100	-0.0379	0.3154	0.0200	0.4560

Table 4: Monte Carlo Results for BLP Instruments

markets	firm size	products per market	median bias	median absolute deviation from α_0	rejection prob. at true α	power of test of $\alpha = 0$
1	2	20	-0.0795	0.3155	0.1510	0.4387
1	2	60	-0.0202	0.1580	0.0893	0.7222
1	2	100	-0.0194	0.1250	0.0836	0.7462
1	10	20	-0.0854	0.3049	0.0794	0.2487
1	10	60	-0.0247	0.1749	0.1130	0.6710
1	10	100	-0.0196	0.1358	0.0762	0.7623
3	varied	20	-0.0241	0.1819	0.0801	0.6286
3	varied	varied	-0.0047	0.0932	0.0441	0.7854
3	varied	60	-0.0090	0.0960	0.0513	0.7678
3	varied	100	-0.0027	0.0760	0.0562	0.8193
3	2	20	-0.0238	0.1766	0.0843	0.6128
3	2	varied	-0.0097	0.0999	0.0592	0.7653
3	2	60	0.0011	0.0930	0.0501	0.7898
3	2	100	0.0003	0.0736	0.0340	0.8338
3	10	20	-0.0262	0.1837	0.1002	0.6092
3	10	varied	-0.0122	0.1000	0.0852	0.7916
3	10	60	-0.0102	0.1007	0.0661	0.7768
3	10	100	-0.0054	0.0767	0.0662	0.8175
20	varied	20	0.0036	0.0703	0.0190	0.7850
20	varied	varied	0.0006	0.0390	0.0593	0.8593
20	varied	60	-0.0004	0.0369	0.0561	0.8509
20	varied	100	-0.0003	0.0287	0.0210	0.9000
20	2	20	-0.0013	0.0685	0.0633	0.7801
20	2	varied	0.0021	0.0402	0.0411	0.8537
20	2	60	0.0035	0.0385	0.0644	0.8632
20	2	100	-0.0003	0.0286	0.0483	0.8813
20	10	20	0.0065	0.0663	0.0220	0.7840
20	10	varied	-0.0008	0.0385	0.0522	0.8554
20	10	60	-0.0023	0.0365	0.0641	0.8707
20	10	100	-0.0027	0.0298	0.0481	0.8826

Table 5: Monte Carlo Results for Cost Instruments

markets	firm size	products per market	median bias	median absolute deviation from α_0	rejection prob. at true α	power of test of $\alpha = 0$
1	2	20	-0.3318	0.6416	0.1054	0.1486
1	2	60	-0.3589	0.6896	0.0842	0.1032
1	2	100	-0.3272	0.6853	0.0874	0.1206
1	10	20	1.4864	28.2989	0.3064	0.3064
1	10	60	-0.3112	0.6440	0.0521	0.0922
1	10	100	-0.3156	0.6748	0.1117	0.1368
3	varied	20	-0.2828	0.6433	0.0130	0.0560
3	varied	varied	-0.3300	0.7105	0.0110	0.0460
3	varied	60	-0.3228	0.7043	0.0090	0.0590
3	varied	100	-0.3146	0.6614	0.0060	0.0470
3	2	20	-0.3583	0.7749	0.0912	0.1082
3	2	varied	-0.3333	0.6597	0.0160	0.0551
3	2	60	-0.3485	0.7714	0.0110	0.0591
3	2	100	-0.3118	0.7599	0.0340	0.0791
3	10	20	-0.3069	0.7160	0.0150	0.0520
3	10	varied	-0.3049	0.7559	0.0090	0.0560
3	10	60	-0.3540	0.7290	0.0120	0.0460
3	10	100	-0.3341	0.7353	0.0250	0.0581
20	varied	20	-0.3111	0.7932	0.0100	0.0620
20	varied	varied	-0.2830	0.7370	0.0090	0.0580
20	varied	60	-0.3471	0.8158	0.0080	0.0450
20	varied	100	-0.3545	0.7563	0.0060	0.0530
20	2	20	-0.3432	0.8074	0.0150	0.0600
20	2	varied	-0.3514	0.7758	0.0130	0.0570
20	2	60	-0.3504	0.8160	0.0060	0.0460
20	2	100	-0.3279	0.8166	0.0080	0.0580
20	10	20	-0.3292	0.7525	0.0100	0.0430
20	10	varied	-0.3570	0.8237	0.0090	0.0500
20	10	60	-0.3387	0.8265	0.1533	0.1814
20	10	100	-0.3454	0.7592	0.0090	0.0470

Table 6: Monte Carlo Results for BLP Instruments with Constant Markups

markets	firm size	products per market	median bias	median absolute deviation from α_0	rejection prob. at true α	power of test of $\alpha = 0$
1	2	20	-0.0614	0.3010	0.1470	0.4673
1	2	60	-0.0148	0.1538	0.0843	0.7329
1	2	100	-0.0185	0.1233	0.0604	0.7613
1	10	20	-0.0493	0.2920	0.0906	0.2867
1	10	60	-0.0231	0.1724	0.0915	0.6915
1	10	100	-0.0192	0.1329	0.0765	0.7533
3	varied	20	-0.0100	0.1694	0.0582	0.6790
3	varied	varied	-0.0085	0.0934	0.0654	0.7827
3	varied	60	-0.0099	0.0969	0.0350	0.7778
3	varied	100	-0.0025	0.0736	0.0431	0.8317
3	2	20	-0.0198	0.1652	0.0512	0.6305
3	2	varied	-0.0078	0.0946	0.0765	0.7736
3	2	60	0.0015	0.0907	0.0240	0.8008
3	2	100	0.0011	0.0722	0.0472	0.8312
3	10	20	-0.0147	0.1765	0.0290	0.6630
3	10	varied	-0.0097	0.0968	0.0361	0.7936
3	10	60	-0.0096	0.0977	0.0701	0.7808
3	10	100	-0.0055	0.0748	0.0531	0.8206
20	varied	20	0.0039	0.0693	0.0731	0.7675
20	varied	varied	-0.0002	0.0371	0.0581	0.8768
20	varied	60	0.0004	0.0363	0.0320	0.8639
20	varied	100	-0.0002	0.0282	0.0230	0.8979
20	2	20	0.0015	0.0643	0.0150	0.7978
20	2	varied	0.0020	0.0382	0.0561	0.8527
20	2	60	0.0040	0.0370	0.0260	0.8619
20	2	100	-0.0006	0.0283	0.0350	0.8919
20	10	20	0.0045	0.0607	0.0711	0.7886
20	10	varied	-0.0000	0.0390	0.0703	0.8574
20	10	60	-0.0021	0.0362	0.0552	0.8645
20	10	100	-0.0023	0.0300	0.0733	0.8855

Table 7: Monte Carlo Results for Cost Instruments with Constant Markups

markets	firm	products per market	$P(BLP\ F > 10)$	$P(F > 10)$	median of $BLP\ F$	median of F
3	varied	20	0.8030	0.7270	12.2745	13.0851
3	varied	varied	0	0.1080	4.6991	5.3294
3	varied	60	0	0.0120	1.8279	2.3794
3	varied	100	0	0.0020	0.9503	1.4793
3	2	20	0	0	0.2324	0.8984
3	2	varied	0	0	0.3586	1.0101
3	2	60	0	0	0.0494	0.7626
3	2	100	0	0	0.0247	0.7411
3	10	20	0.0340	0.0370	1.4653	2.1550
3	10	varied	0.9930	0.8400	13.9655	14.8231
3	10	60	0	0	0.4476	0.9579
3	10	100	0	0	0.2314	0.8708
20	varied	20	1.0000	1.0000	79.9612	80.7936
20	varied	varied	0.9980	0.7080	12.3633	12.9017
20	varied	60	0.9910	0.6920	11.8897	12.1792
20	varied	100	0	0.2230	6.1283	6.7278
20	2	20	0	0.0100	1.7318	2.1817
20	2	varied	0	0.0240	2.4032	2.9314
20	2	60	0	0	0.3260	0.8944
20	2	100	0	0	0.1686	0.8750
20	10	20	0.6470	0.6380	12.1892	12.3467
20	10	varied	1.0000	1.0000	93.5960	93.5900
20	10	60	0	0.0210	3.0468	3.4386
20	10	100	0	0.0020	1.5715	2.0943

Table 8: Monte Carlo Results for First Stage Tests