Ownership Consolidation and Product Characteristics:

A Study of the U.S. Daily Newspaper Market

Online Appendix

Ying Fan
Department of Economics, University of Michigan

C Invertibility of the Penetration Function

In this appendix, I show that the invertibility result in BLP can be extended to a multiple discrete choice model. I only show the extension for a model where the number of products that an individual can buy is limited to at most two. The result can be easily extended to a model in which consumers can choose up to $\bar{n} \leq J$ products, where J is the total number of products available in a choice set.

Penetration Function

Let $\Phi(\cdot)$ represent the distribution function of the random term ς_i. The penetration function in Section 2.1 is given by

$$\begin{aligned} \mathscr{s}_j(\boldsymbol{\delta}, \boldsymbol{x}; \boldsymbol{\sigma}, \kappa) &= \int \Psi_j^{(1)}(\boldsymbol{\delta}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma})\, d\Phi(\varsigma_i) \\ &\quad + \sum_{j' \neq j} \int \left(\Psi_{j,j'}^{(2)}(\boldsymbol{\delta}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa) - \Psi_j^{(3)}(\boldsymbol{\delta}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa) \right) d\Phi(\varsigma_i), \end{aligned}$$

where

$$\Psi_j^{(1)}(\boldsymbol{\delta}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}) = \frac{\exp(\delta_j + \vartheta_{ij})}{1 + \sum_{h=1}^{J} \exp(\delta_h + \vartheta_{ih})},$$

is the probability that newspaper j is chosen as the first newspaper (ϑ_{ij} is the deviation of household i's utility from the mean utility), and the probability that newspaper j is chosen as the second newspaper when j' is the first best is given by the difference between the followings:

$$\begin{aligned} \Psi_{j,j'}^{(2)}(\boldsymbol{\delta}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa) &= \frac{\exp(\delta_j + \vartheta_{ij})}{\exp(\kappa) + \sum_{h \neq j'} \exp(\delta_h + \vartheta_{ih})}, \\ \Psi_j^{(3)}(\boldsymbol{\delta}, \boldsymbol{x}, \varsigma_i; \boldsymbol{\sigma}, \kappa) &= \frac{\exp(\delta_j + \vartheta_{ij})}{\exp(\kappa) + \sum_{h=1}^{J} \exp(\delta_h + \vartheta_{ih})}. \end{aligned}$$

Invertibility

Since all statements in this section are true for any given $(\boldsymbol{x}, \boldsymbol{\sigma}, \kappa)$, these arguments in $\mathscr{s}_j$ are omitted for expositional simplicity.

The proof of the invertibility result is slightly different from that in BLP. BLP define a function $F : R^J \to R^J$ pointwise as $F_j(\boldsymbol{\delta}) = \delta_j + \ln s_j - \ln \mathscr{s}_j(\boldsymbol{\delta})$ and show that F is a contraction when an upper bound on the value taken by F is imposed. For a single discrete choice model, the value of δ_j that solves $\sum_{h=1}^{J} s_h = \sum_{h=1}^{J} \mathscr{s}_h(\boldsymbol{\delta})$ when $\delta_{j'} = -\infty$ for $\forall j' \neq j$ is the upper bound of the jth dimension of a fixed point of F. In a multiple discrete choice model, however, this value does not exist when $\left(\sum_{h=1}^{J} s_h\right)$ is larger than 1.[24]

I first prove the existence and uniqueness of the solution to $\mathscr{s}_j(\boldsymbol{\delta}, \boldsymbol{x}; \boldsymbol{\sigma}, \kappa) = s_j$ for all j directly without using the function F. I then verify that all conditions in BLP hold so that F is indeed a contraction mapping – when an upper bound is imposed.

The following inequalities, which will be proven at the end of this section, are useful in the proof:

$$\partial \mathscr{s}_j / \partial \delta_j < \mathscr{s}_j \tag{C.4}$$

$$\partial \mathscr{s}_j / \partial \delta_j > 0 \tag{C.5}$$

$$\partial \mathscr{s}_j / \partial \delta_h < 0 \text{ when } h \neq j \tag{C.6}$$

$$\sum_{h=1}^{J} \left(\partial \mathscr{s}_j / \partial \delta_h\right) > 0 \tag{C.7}$$

Inequalities (C.5), (C.6) and (C.7) imply that the Jacobian of $\mathscr{s}$ has a dominant diagonal. Therefore, there is a unique solution to the equation system of $\mathscr{s}_j(\boldsymbol{\delta}) = s_j$ for all j.[25]

I now prove that all conditions in the theorem in BLP hold.

Condition (1): Inequalities (C.4) and (C.6) imply that

$$\begin{aligned} \partial F_j(\boldsymbol{\delta}) / \partial \delta_j &= 1 - \left(\partial \mathscr{s}_j / \partial \delta_j\right) / \mathscr{s}_j > 0 \\ \partial F_j(\boldsymbol{\delta}) / \partial \delta_h &= -\left(\partial \mathscr{s}_j / \partial \delta_h\right) / \mathscr{s}_j > 0 \text{ when } h \neq j. \end{aligned}$$

Also, inequality (C.7) implies that

$$\sum_{h=1}^{J} \partial F_j(\boldsymbol{\delta}) / \partial \delta_h = 1 - \sum_{h=1}^{J} \left(\partial \mathscr{s}_j / \partial \delta_h\right) / \mathscr{s}_j < 1.$$

[24] In a single discrete choice model, $\sum_{h=1}^{J} s_h < 1$, while in a multiple discrete choice model, the sum of market penetration for all products $\sum_{h=1}^{J} s_h$ can be larger than 1. But the supremum of $\sum_{h=1}^{J} \mathscr{s}_h(\boldsymbol{\delta})$ is 1 when $\delta_{j'} = -\infty$ for $\forall j' \neq j$.

[25] See McKenzie, Lionel (1959), "Matrices with dominant diagonals and economic theory." In *Mathematical methods in the social sciences* (Kenneth Joseph Arrow, Samuel Karlin, and Patrick Suppes, eds.), 47-62, Stanford University Press.

Condition (2): Given the monotonicity of F in all dimensions of $\boldsymbol{\delta}$, a lower bound of function F is $\underline{\delta} = \min_j \left(\lim_{\boldsymbol{\delta}\to-\infty^J} F_j(\boldsymbol{\delta})\right)$.

Condition (3): I have already shown that the equation system of $\mathscr{s}_j(\boldsymbol{\delta}) = s_j$ has a unique solution. This implies that the mapping F has a unique fixed point. Denote the fixed point by $\boldsymbol{\delta}^*$. Then, $F_j(\boldsymbol{\delta}^*) = \delta_j^*$ for all j. Note that $F_j(\boldsymbol{\delta}^* + \Delta) - \left(\delta_j^* + \Delta\right) = \ln s_j - \ln \mathscr{s}_j(\boldsymbol{\delta}^* + \Delta)$ is strictly decreasing in Δ as implied by inequality (C.7). Therefore, $F_j(\boldsymbol{\delta}^* + \Delta) < \left(\delta_j^* + \Delta\right)$ for any $\Delta > 0$. Define $\bar{\delta}_j = \delta_j^* + \Delta$. Then, $F_j\left(\bar{\boldsymbol{\delta}}\right) < \bar{\delta}_j$ for any j. By inequality (C.6), $F_j(\boldsymbol{\delta}) < \delta_j$ for any $\boldsymbol{\delta}$ such that $\delta_j = \bar{\delta}_j$ and $\delta_{j'} \le \bar{\delta}_{j'}$ for all j'.

I now show inequalities (C.4) to (C.7). Three observations are important:

$$0 < \Psi_j^{(1)}, \Psi_{j,j'}^{(2)}, \Psi_j^{(3)} < 1;\ \Psi_{j,j'}^{(2)} > \Psi_j^{(3)};\ \Psi_j^{(1)} > \Psi_j^{(3)}.$$

Inequalities (C.4) and (C.6) follow directly from the three observations:

$$\begin{aligned}
\partial \mathscr{s}_j / \partial \delta_j &= \int \Psi_j^{(1)}\left(1 - \Psi_j^{(1)}\right) d\Phi(\varsigma) + \sum_{j' \neq j} \int \left[\Psi_{j,j'}^{(2)}\left(1 - \Psi_{j,j'}^{(2)}\right) - \Psi_j^{(3)}\left(1 - \Psi_j^{(3)}\right)\right] d\Phi(\varsigma) \\
&< \int \Psi_j^{(1)} d\Phi(\varsigma) + \sum_{j' \neq j} \int \left(\Psi_{j,j'}^{(2)} - \Psi_j^{(3)}\right) d\Phi(\varsigma) = \mathscr{s}_j, \\
\partial \mathscr{s}_j / \partial \delta_h &= -\int \Psi_j^{(1)} \Psi_h^{(1)} d\Phi(\varsigma) + \int \Psi_j^{(3)} \Psi_h^{(3)} d\Phi(\varsigma) + \sum_{j' \neq j,h} \int \left(-\Psi_{j,j'}^{(2)} \Psi_{h,j'}^{(2)} + \Psi_j^{(3)} \Psi_h^{(3)}\right) d\Phi(\varsigma) \\
&< \sum_{j' \neq j,h} \int \left(-\Psi_{j,j'}^{(2)} \Psi_{h,j'}^{(2)} + \Psi_j^{(3)} \Psi_h^{(3)}\right) d\Phi(\varsigma) < 0 \text{ when } h \neq j.
\end{aligned}$$

To show inequality (C.7), note that $\sum_{h=1}^J \frac{\partial \mathscr{s}_j(\boldsymbol{\delta})}{\partial \delta_h} = \frac{\partial \mathscr{s}_j(\boldsymbol{\delta}+\Delta)}{\partial \Delta}|_{\Delta=0}$, and

$$\begin{aligned}
&\frac{\partial \mathscr{s}_j(\boldsymbol{\delta} + \Delta)}{\partial \Delta}|_{\Delta=0} \\
&= \int \left(\Psi_j^{(1)}\right)^2 \frac{1}{e^{\delta_j + \vartheta_{ij}}} d\Phi(\varsigma) + \sum_{j' \neq j,0} \int \left[\left(\Psi_{j,j'}^{(2)}\right)^2 - \left(\Psi_j^{(3)}\right)^2\right] \frac{e^{\kappa}}{e^{\delta_j + \vartheta_{ij}}} d\Phi(\varsigma) > 0.
\end{aligned}$$

Combining inequalities (C.6) and (C.7) yields inequality (C.5).

D Instrumental Variables

The "excluded" instrumental variables that are assumed uncorrelated with the demand error ξ_{jct}:

- BLP-style instrument (IV_1)
  - the number of competitors
- cost shifters (IV_2, IV_3, IV_4)
  - frequency of publication
  - households in the home county of newspaper j (It is correlated with Q_{jt} in the average cost of printing and delivery in (10).)
  - households in the home counties of other papers that are close-by and are of the same owner as newspaper j (It is correlated with Q_{jt} in (10).)
- demographics in counties of newspaper j's market (excluding county c) (IV_5, IV_6)
  - weighted average education and median age (weighted by households in a county)
- demographics in the counties of newspaper j's competitors (excluding counties in j's market) ($IV_7, IV_8, IV_9, IV_{10}$)
  - weighted average education, median income, median age and urbanization (weighted by households in a county)

The table below reports the results of the first-stage regression. Standard errors are clustered by newspaper. The estimates are largely significant.

First-stage Regression Result

	newshole[a]		opinion		reporter		local news ratio		variety		price	
	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.
Included IV												
log(households in the market)	0.169**	0.035	0.449**	0.264	-0.145	1.185	0.331**	0.053	0.129**	0.028	12.804**	2.653
morning edition	0.052**	0.026	0.654**	0.313	0.311	0.996	-0.070	0.070	-0.092**	0.039	2.701	2.525
local dummy	-0.008	0.027	-0.041	0.217	0.236	0.863	0.007	0.038	-0.002	0.020	1.044	2.539
county distance (1000km)	-0.074	0.366	-0.339	3.319	8.000	12.139	1.373**	0.539	0.341**	0.205	103.969**	43.441
education	-0.074**	0.032	-1.520**	0.371	-11.998**	1.418	-0.436**	0.068	0.227**	0.038	15.894**	2.170
median income ($10000)	0.215*	0.154	1.343*	1.020	1.953	3.742	0.210	0.334	0.163	0.234	38.557**	14.035
median age	-0.005**	0.002	-0.013	0.013	-0.021	0.049	0.004	0.004	0.003	0.002	0.028	0.160
urbanization	0.076	0.293	2.071	1.713	9.786*	7.018	0.077	0.560	0.161	0.362	82.087**	25.110
time	-0.006	0.005	-0.003	0.041	-0.020	0.180	-0.005	0.010	0.002	0.006	-0.422	0.440
Excluded IV												
IV_1	0.034**	0.018	0.235*	0.150	1.084**	0.457	-0.029**	0.008	-0.012**	0.004	-1.871**	0.716
IV_2	3.383**	0.356	-5.718*	3.745	-0.295	14.841	4.867**	0.895	2.679**	0.535	229.767**	35.724
IV_3	3.176**	0.563	33.451**	3.265	171.795**	17.737	-1.461**	0.536	-0.536**	0.229	-55.646**	22.626
IV_4	0.107	0.478	4.334	6.596	16.415	26.662	-0.507	0.482	-0.314	0.334	-16.786	39.376
IV_5	0.724**	0.328	7.210**	2.914	29.476**	10.128	0.089	0.438	-0.466**	0.156	71.055**	26.044
IV_6	-0.002	0.001	-0.019*	0.012	-0.081**	0.044	0.002	0.003	0.004**	0.001	-0.082	0.114
IV_7	-0.659**	0.368	-7.136**	3.306	5.685	11.343	0.737	0.732	-0.349	0.399	-41.265	36.571
IV_8	0.241	0.356	5.502**	3.251	13.255*	10.183	-0.384	0.701	0.474*	0.340	41.246*	27.609
IV_9	-0.007**	0.003	-0.031*	0.023	-0.121*	0.083	-0.005	0.005	-0.003*	0.002	0.416**	0.235
IV_{10}	-0.091	0.120	-1.079	1.071	-13.159**	3.920	0.284	0.244	0.002	0.096	-16.563*	11.301
F test												
F(19, 943)	144.10		44.91		46.14		32.67		11.02		48.06	
F test of excluded IV												
F(10, 943)	16.64		26.72		28.82		9.19		6.92		11.43	

** indicates 95% level of significance. * indicates 90% level of significance.

[a] The news hole is $n_{jt} - a\left(r_{jt}, q_{jt}, H_{jt}; \hat{\eta}, \hat{\boldsymbol{\lambda}}\right)$, where $\hat{\eta}$ and $\hat{\boldsymbol{\lambda}}$ are the estimated advertising demand parameters reported in Table 4.

The instrumental variables that are assumed uncorrelated with the advertising demand error ι_{jt} in (8) include weighted average education, median income as well as median age and urbanization in newspaper j's market (weighted by households in a county).

The instrumental variables that are assumed uncorrelated with the supply-side error terms ζ_{jt}, ω_{jt}, ν_{kjt} in the estimation equations (16), (17) and (18) include weighted average education, median income, median age and urbanization in newspaper j's market (weighted by households in a county), frequency of publication, the number of households in newspaper j's market as well as households in the home counties of other papers that are close-by and are of the same owner as newspaper j.

E Details on GMM

To define the GMM objective function, I solve for all error terms and express them as functions of the data and the model parameters. Specifically,

$$\xi_{jct} = \delta_{jct}\left(\boldsymbol{s}_{ct}; \boldsymbol{\sigma}, \kappa\right) - p_{jt}\alpha - \boldsymbol{x}_{jt}\boldsymbol{\beta} - \boldsymbol{x}_{jt}\boldsymbol{z}_{ct}\boldsymbol{\theta} - \boldsymbol{y}_{jct}\boldsymbol{\psi} - \boldsymbol{z}_{ct}\boldsymbol{\varphi} + (t - t_0)\,\rho \tag{E.8}$$

$$\iota_{jt} = \log a_{jt} - \eta - \lambda_0 \log H_{jt} - \lambda_1 \log q_{jt} - \lambda_2 \log r_{jt} \tag{E.9}$$

$$\omega_{jt} = ac_{jt}^{(q)} - (\gamma_1 + \gamma_2 f_{jt} + \gamma_3\left(x_{1jt} + a_{jt}\right)) log(Q_{jt})^{\gamma_4} \tag{E.10}$$

$$\zeta_{jt} = r_{jt} - \bar{\zeta} - \frac{\gamma_3}{1 + 1/\lambda_2} log(Q_{jt})^{\gamma_4} q_{jt} \tag{E.11}$$

$$\nu_{kjt} = \sum_{h \in \mathcal{J}_{mt}} \left(\frac{\partial \pi_{ht}^{\mathrm{II}}}{\partial x_{kjt}} + \sum_{j' \in \mathcal{J}_{g(jt)}} \frac{\partial \pi_{ht}^{\mathrm{II}}}{\partial p_{j't}} \frac{\partial p_{j't}^{*}}{\partial x_{kjt}} \right) - \tau_{k0} - \tau_{k1} x_{kjt}. \tag{E.12}$$

In (E.10), $ac_{jt}^{(q)}$ is solved from the first-order condition with respect to the subscription price (equation (17)). In (E.12), the partial derivative $\frac{\partial p_{j't}^{*}}{\partial x_{kjt}}$ is computed following the approach described at the end of Section 2.

These error terms are plugged into the moment conditions that they are uncorrelated with the instrumental variables explained in Appendix D. Let Θ be the collection of all model parameters. Denote the empirical moment conditions by $g(\Theta)$. Then, the GMM objective function is $g(\Theta)' W g(\Theta)$, where W is a weighting matrix. I first take the efficient weighting matrix assuming all error terms are homoscedastic to compute an estimate. This estimate is used to obtain a consistent estimate of the covariance matrix, the inverse of which is used as the weighting matrix to obtain another estimate. This process is iterated once more as the estimate does not change much any more from the last iteration.

F Estimation Results when 80% of Total Circulation is Used as a Criterion for Defining the Market of a Newspaper

		parameter	estimate	s.e.
Utility	price ($100)	α	-1.400**	0.276
	log(1+x_1)	β_1	0.825**	0.176
	log(1+x_2)	β_2	-0.003	0.035
	log(1+x_3)	β_3	0.034*	0.023
	weight on newshole in x_1	1		
	weight on opinion in x_1	ϖ_2	0.765	3.259
	weight on reporters in x_1	ϖ_3	4.855*	2.956
	log(1+x_2), education	θ_1	-0.021*	0.014
	log(1+x_2), median age	θ_2	0.026	0.049
	log(1+x_2), s.d.	σ	0.056	0.151
	log(households in the market)	ψ_1	-0.852**	0.149
	morning edition	ψ_2	0.099*	0.076
	local dummy	ψ_3	0.328**	0.083
	county distance (1000km)	ψ_4	-3.140**	1.055
	constant	φ_0	6.071**	1.226
	education	φ_1	0.881**	0.376
	median income ($10000)	φ_2	-0.064	0.376
	median age	φ_3	0.044**	0.008
	urbanization	φ_4	0.755**	0.174
	time	ρ	0.141**	0.023
	diminishing utility	κ	1.896**	0.744
Display ad demand	ad market size	λ_0	0.045	0.113
	total circulation	λ_1	1.673**	0.131
	ad rate	λ_2	-1.194**	0.206
	constant	ϕ	-1.626	1.538
Avg cost of circulation	constant	γ_1	-49.357	1031.900
	frequency	γ_2	1.751**	0.669
	pages in a year	γ_3	0.022	0.019
	economies of scale/scope	γ_4	-0.442**	0.109
Marginal cost of ad sales		$\bar{\zeta}$	12.960**	2.429
Slope of the fixed cost for x_1	constant	τ_{10}	99299**	34481
	x_1	τ_{11}	0.00799	138.720
Slope of the fixed cost for x_2	constant	τ_{20}	89.252	1675.200
	x_2	τ_{21}	11.962	30.011
Slope of the fixed cost for x_3	constant	τ_{30}	0.0001	4057.000
	x_3	τ_{31}	0.103	157.610
Preprint profit	circulation	μ_1	152.460	361.440
	square of circulation	μ_2	-0.002*	0.001

** indicates 95% level of significance.

* indicates 90% level of significance.