Spatial Differentiation and Price Discrimination in the Cement Industry: Evidence from a Structural Model

Online Appendix

Nathan H. Miller
Georgetown University

Matthew Osborne
University of Toronto

The asymptotic properties of the estimator rest on a number of assumptions, some of which can be evaluated numerically. First, point identification can fail if multiple candidate parameters produce equilibrium predictions that are identical once aggregated to the level of the available data. This is more likely when the data are relatively coarse so that aggregation entails a substantial loss of information.

We conduct an artificial data experiment to check for this sort of aggregation problem in our empirical application. We pair a vector of "true" parameters with 40 randomly-drawn sets of exogenous data. Both the parameters and the data are chosen to mimic the application. For each of set of exogenous data, we compute equilibrium, generate the relevant aggregated data, and estimate the model. We argue that the parameters are reasonably identified if the estimates are close to the true parameters.[1]

Table 1 shows the results of the artificial data experiment. Interpretation is complicated somewhat because we use non-linear transformations to constrain the some of coefficients (e.g., $\beta^p < 0$) as discussed in Appendix C in the paper. Nonetheless, it is clear that the means of the estimated coefficients are close to transformed true parameters. The means of the price and distance coefficients are within 6 percent and 11 percent of the truth, respectively. This

[1]The exogenous data includes the plant capacities, the potential demand of counties, the diesel price, the import price, and two cost shifters. We randomly draw capacity and potential demand from the data (with replacement), and we draw the remaining data from normal distributions. Specifically, we use the following distributions: diesel price $\sim N(1, 0.28)$, import price $\sim N(50, 9)$, cost shifter $1 \sim N(60, 15)$, and cost shifter $2 \sim N(9, 2)$. We redraw data that are below zero and data that lead the estimator to nonsensical areas of parameter space. Throughout, we hold plant and county locations fixed to maintain tractability, and rely on the random draws of capacity, potential demand, and diesel prices to create variation in the distances between production capacity and consumers. Each artificial data set includes 21 draws on the exogenous data, with each draw representing a single time-series observation.

Table 1: Artificial Data Test for Identification

Variable	Parameter	Truth (θ)	Transformed ($\widetilde{\theta}$)	Mean Est	RMSE
Demand					
Cement Price	β^p	-0.07	-2.66	-2.51	0.66
Miles×Diesel Price	β^d	-25.00	3.22	2.86	0.59
Import Dummy	β^i	-4.00	-4.00	-6.07	1.23
Intercept	β^c	2.00	2.00	1.11	0.51
Inclusive Value	λ	0.09	-2.31	-1.73	0.54
Marginal Costs					
Cost Shifter 1	α_1	0.70	-0.36	-0.88	0.51
Cost Shifter 2	α_2	3.00	1.10	0.54	0.45
Utilization Threshold	ν	0.90	2.19	1.71	0.59
Over-Utilization Cost	γ	300.00	5.70	6.14	1.05

Notes: Results of estimation on 40 data sets that are randomly drawn based on the "true" parameters listed. The parameters are transformed prior to estimation to place constraints on the parameter signs/magnitudes. Mean Est and RMSE are the mean of the estimated (transformed) parameters and the root mean-squared error, respectively.

precision is notable because the ration of price and distance coefficients determines the unit transportation cost and thereby the degree of spatial differentiation. The other means of the estimated demand coefficients are somewhat farther from the truth. Among the marginal cost parameters, the mean estimated coefficients are accurate for the utilization threshold and the over-utilization cost but less accurate for the constant cost shifters. We conclude that the primary coefficients of interest (for spatial considerations) are likely well-identified but that some skepticism of the other coefficients may be appropriate, especially with regard to the constant marginal cost shifters.

Second, the continuity and differentiability of the implicit solution to the firms' first order condition fails if multiple equilibria are present. We search for only a single equilibria in the inner loop in our application. For robustness, we conduct a Monte Carlo experiment and search for the existence of multiple equilibria. In particular, we compute equilibrium at eleven different starting points for thousands of randomly-drawn candidate parameter vectors. We then evaluate whether, for each given candidate parameter vector, the computed equilibrium prices are sensitive to the starting points.[2] More precisely, for each candidate

[2]We consider 300 parameter vectors for each of the 21 years in the sample, for a total of 6,300 candidate parameter vectors. For each $\theta_i \in \boldsymbol{\theta}$, we draw from the distribution $N(\widehat{\mu}_i, \widehat{\sigma}_i^2)$, where $\widehat{\mu}_i$ and $\widehat{\sigma}_i$ are the coefficient and standard error, respectively, reported in mo2011. We then compute the numerical equilibrium

parameter vector, we calculate the standard deviation of each equilibrium price across the eleven starting points. (So there are 1,260 standard deviations for a typical equilibrium price vector of 1,260 plant-area elements.) The results indicate that the maximum standard deviation, over all candidate parameter vectors and all plant-area prices, is zero to computer precision. Thus, the Monte Carlo experiment finds no evidence of multiple equilibria. This may be unsurprising because, theoretically, uniqueness is ensured for two close cousins of our model: nested logit demand with single-plant firms (Mizuno 2003) and logit demand with increasing marginal costs and multi-plant firms (Konovalov and Sándor 2010).

for each parameter vector, using eleven different starting vectors. We define the elements of the starting vectors to be $p_{jnt} = \phi \overline{p}_t$, where $\overline{p}_t$ is the average price of portland cement and $\phi = 0.5, 0.6, \ldots, 1.4, 1.5$. Thus, we start the equation solver at initial prices that are sometimes larger and sometimes smaller than the average prices observed in the data. The equal-solver computes numerical equilibria for 90 percent of the candidate vectors. See Appendix C for a discussion of non-convergence in the inner-loop.