Panel Data Reduces Bias in Entry Models

Allan Collard-Wexler Economics Department, Northwestern University

Abstract

Entry Models such as Bresnahan and Reiss(91)[4] can underestimate the effect of competition. If the profitability of markets is mismeasured, this introduces an positive correlation between unoberserved profitability and the number of firms in a market. Using data on entry and exit patterns in the Ready-Mix Concrete Industry from $1976-1999$, I show that using fixed effects in a Bresnahan-Reiss entry model reduces the coefficient on demand by 50% and increases the coefficient on competition by 100% compared to the no fixed-effect benchmark.

1 Introduction

Following the work of Bresnahan and Reiss(91)[4], entry models have become popular for empirical work in industrial organization. This is due to their solid grounding in economic theory, in which they are an outcome of a well specified game between firms, and to the ease of estimating a Bresnahan-Reiss model versus other more sophisticated models of entry such as Aguiregabierra and Mira $(04)[1]$ or Bajari-Benkard-Levin $(05)[3]$. Along with popularity of entry models come the use of these model to increasingly less ideal data, in which markets may be ill-defined, sources of actual demand poorly measured and differences between competitors unrecognized. Following the lead of work in labor economics on estimating the returns to education (Angrist and Krueger (1999)[2]), panel data is used to identify and reduce bias in entry models. Empirical estimates show that including market fixed effects raises the coefficient on competition by over 100% and lowers that of demand by 50%, in an application to entry and reaction to demand shocks in the ready-mix concrete industry.

The ready-mix concrete industry is particularly well suited to study entry in geographically segmented markets because it sells such a perishable product: "ASTM C 94 also requires that concrete be delivered and discharged within 1 1/2 hours or before the drum has revolved 300 times after introduction of water to the cement and aggregates".¹ Moreover, concrete is very cheap for its weight. Indeed, one ready-mix producer describes the economics of transportation costs in the ready-mix industry as:

A truckload of concrete contains about 7 cubic yards of concrete. A cubic yard of concrete weights about 4000 pounds and will cost you around 60\$ delivered to your door. That's 1.5 cents a pound. If you go to your local hardware store, you get a bag of manure weighing 10 pounds for 5\$. This means that concrete is cheaper than shit.²

So in practise, the delivery radius of ready-mix is even more constrained and in most cases a ready-mix truck will drive about 20 minutes to deliver its load. This is the cause of concrete's most salient feature: very segmented markets with a limited number of incumbents in each area. Beyond the issue of segmentation, Ready-Mix concrete is a fairly homogenous good. While it is possible to produce several hundred types of Ready-Mix concrete, these mixtures basically use the same ingredients and machinery. As well, in part because of fairly aggressive vertical antitrust policy conducted by the United States Government, the average ready-mix producer is a single plant operator, reducing problems of multiplant decision making.

A ready-mix concrete plant is expensive to set up. The most important cost is constructing the physical structure for loading ready-mix concrete and dispensing it into mixing trucks, as well as other buildings such as garages. In addition, the land on which the plant is located much be purchased, as well as ready-mix trucks which cost around US\$120 000 new and US\$ 40 000 used. All together, the cost of setting up a readymix plant is between 3 and 4 millions dollars based, on a dozen interviews of established producers building new plants, and the average value of an existing plant is about 2 million dollars in 1997.

There are few large costs for shutting down a ready-mix plant. Trucks can be sold on a competitive used vehicle market, and land can be resold

¹See p.96 of Kosmatka, Kerkhoff et al. $(02)[8]$.

²Phone interview, January 2005.

for other uses. However, the plant itself is a total loss. At best it can be resold for scrap metal, but many ready-mix plants are left on site because the cost of dismantling them outweights any benefits. Over all, the scrap value of a plant is at most 1 millions dollars, based almost exclusively on the value of a ready-mix truck fleet and the value of the land on which the plant is located.

2 Statics: Bresnahan Reiss Models

Static models, such as Bresnahan-Reiss[4], can be used to investigate the presence of sunk costs in ready-mix concrete. This model does not compute the value function from period profits. Instead, the value function is directly estimated, without reference to what will happen in the future, from the current configuration of firms in a market. This "reduced form" model is used to check for a wide variety of empirical problems, such as different assumptions on the shocks to firms profits. Being a reduced form, it cannot be used for many counterfactual exercices.

The first Bresnahan Reiss[4] model is based on two assumptions:

1. Firms that Enter make Positive Profits

 $\pi(N, X_m) + \varepsilon_m > 0$

2. If an extra firm entered it would make negative profits:

$$
\pi(N+1, X_m) + \varepsilon_m < 0
$$

where $\pi(N, X_m)$ is the oberservable component of profit depending on demand side factors X_m and the number of symmettric competitors in a market N, while ε_m are unobserved components of profitability common to all firms in a market.

Assume market level shocks ε_m have a normal distribution with zero mean and unit variance. The probability of observing a market X_m with N plants is:

$$
\Phi[-\pi(N+1, X_m)] - \Phi[-\pi(N, X_m)]1(N > 0)
$$

where $\Phi(.)$ is the cumulative distribution function of the standard normal. Parameters of this model can be estimated via Maximum Likelihood.

Firms make sunk, unrecoverable investments when they enter a market. The decision of an incumbent firm to remain in a market differs from the decision of an entrant to build a new plant. The next series of models deal with this difference.

Figure 1: Entry Threshold ψ and Exit Threshold ϕ based on static profits.

2.1 Bresnahan-Reiss Model of Exit

The Bresnahan-Reiss[5] Model of Exit distinguishes between two types of firms: firms which are already active and firms which are deciding to enter the market. Entrants and incumbents have the same shocks and profits, and hence the same continuation values. However, entrants always have lower profits than incumbents, since they pay an entry cost that incumbents do not, as is shown by figure 1. This implies that there cannot be simultaneous entry and exit: either firms exit, enter, or nothing happens. This is a feature of all models which do not have firm specific shocks and where firms are symmettric: they cannot rationalize the same type of plant in the same market making different choices. Thus market-years in which there is both entry and exit are dropped. For this paper, yearly data is used for markets with on average less than 3 incumbents. Thus less than 5% of markets need to be dropped. Moreover, including these markets does not significantly change estimated parameters. Three regimes need to be considered: entry, exit and stasis.

1. Net Entry: $N_t > N_{t-1}$

$$
\pi(N_t, X_{mt}) + \varepsilon_{mt} > \psi
$$

$$
\pi(N_t + 1, X_{mt}) + \varepsilon_{mt} < \psi
$$

2. Net Exit: $N_t < N_{t-1}$

$$
\pi(N_t, X_{mt}) + \varepsilon_{mt} > \phi
$$

$$
\pi(N_t + 1, X_{mt}) + \varepsilon_{mt} < \phi
$$

3. No Net Change: $N_t = N_{t-1}$

$$
\pi(N_t, X_{mt}) + \varepsilon_{mt} > \phi
$$

$$
\pi(N_t + 1, X_{mt}) + \varepsilon_{mt} < \psi
$$

where ϕ is the entry fee that an existing firm pays to enter the market and ψ is the scrappage value of a firm. Entry fees and scrap value are not identified from fixed costs, since it is always possible to increase fixed costs and entry/exit fees by the same amount without changing the probability a market configuration. Yet, the difference between entry and exit fees is identified and can be compared to other quantities such as the effect of an extra competitor.

These equations can be combined into:

$$
\pi(N_t, X_{mt}) + \varepsilon_{mt} > 1(N_t > N_{t-1})\psi + 1(N_t \le N_{t-1})\phi \tag{1}
$$

$$
\pi(N_t + 1, X_{mt}) + \varepsilon_{mt} < 1(N_t \ge N_{t-1})\psi + 1(N_t < N_{t-1})\phi \tag{2}
$$

The probability of observing a market X_m with N_t plants today and N_{t-1} plants in the last period is:

$$
\Phi[-\pi(N_t+1, X_{mt}) + 1(N_t+1 \ge N_{t-1})\psi + 1(N_t+1 < N_{t-1})\phi] \n- \Phi[-\pi(N_t, X_{mt}) + 1(N_t > N_{t-1})\psi + 1(N_t \le N_{t-1})\phi]1(N_t > 0)
$$

which is used to form a maximum likelihood estimator.

The assumption that the epsillon's are serially uncorrelated within markets is heroic. Characteristics of the market that are not observed in the first period, such as a vast road network requiring a large amount of concrete, are the same in each subsequent period. Serial correlation of ε only affect the standard errors of maximum likelihood. However, this pattern of correlation can be used to indentify bias in the Bresnahan-Reiss model. In the next section we will discuss the impact of unmeasured components of profitability on estimated coefficients.

2.2 Unobserved Profitability

The canonical entry model will estimate the profit function for a firm in different markets:

$$
\pi_{it} = X_{it}\beta + g(N_{it}) + \varepsilon_{it} \tag{3}
$$

where ε_{it} is a mean-zero stochastic term which is uncorrelated with both demand (X_{it}) and number of firms (N_{it}) , and $g(.)$ is decreasing. The assumption that ε_{it} is uncorrelated with regressors is frequently violated in the context of entry models. The econometrician may not observe certain components of profitability, but firms most certainly do. They will react by entering in greater numbers in more profitable markets, leading to a positive correlation between ε and N. Likewise, suppose demand in large markets is qualitatively different than in small markets. For instance, multistory buildings are constructed in greater proportion in large markets relative to small markets. This induces an artificial correlation between market size and consumption of concrete.

Unobserved profitability can be decomposed into its correlated components:

$$
\varepsilon_{it} = \delta X_{it} \text{(observed demand)} + \gamma N_{it} \text{(firms)} + \zeta_{it} \tag{4}
$$

where ζ_{it} is an uncorrelated shock, mean zero shock.

If measured and unmeasured demand are positively correlated, say because areas with large numbers of construction workers and projects also have other features which make demand high, then $\delta > 0$. Similarly, if firms react to unmeasured demand shocks by entering, we expect $\gamma > 0$. Note that both statements refer to the correlation between ξ (unobserved demand) and X_{it} or N_{it} , while the values of δ or γ are related to the conditional correlation $E(\varepsilon X|N)$ or $E(\varepsilon N|X)$, which are difficult to make statements about. In the case where the conditional correlation has the same sign as the unconditional correlation, it is possible to sign the bias in this model:

The ordered probit model can be expressed as:

$$
X_{it}\beta + \varepsilon_{it} > -g(N_{it})\tag{5}
$$
\n
$$
X_{it}\beta + \varepsilon_{it} < -g(N_{it} + 1)
$$

Using the expression 4, the ordered probit inequalities become:

$$
X_{it}(\beta + \delta) + \zeta_{it} > -g(N_{it}) - \gamma N_{it}
$$

$$
X_{it}(\beta + \delta) + \zeta_{it} < -g(N_{it} + 1) - \gamma N_{it}
$$

The estimated demand coefficient $(\beta + \delta)$ will be biased upward. Likewise, since the effect of competition is negative, the competitive effects of entry $-[g(N) + \gamma]$ will be biased downwards. If fact, this is what is found in empirical estimates in table 2.

2.3 Bias in Threshold Ratios

The ratio of entry thresholds, used by Bresnahan and Reiss [4] to compute the number of Örms required in a market for its behaviour to be competitive, may be overestimated.

To see this, sign the inequality between true and estimated threshold ratio needs to be evaluated:

$$
\frac{g(N_{it}+1) + \gamma(N_{it}+1)}{g(N_{it}) + \gamma N_{it}} >< \frac{g(N_{it}+1)}{g(N_{it})}
$$

Mutiply both sides of this inequality by $[g(N_{it}) + \gamma N_{it}]g(N_{it})$. This quantity will be a positive number under the assumption that $-g(N_{it} +$ $1) > \gamma N_{it}$, that unobserved components of profitability correlated with the number of firms does not outweight the observed effect. Rearranging and cancelling gives:

$$
\frac{[N_{it}+1]}{N_{it}} \leq \frac{g(N_{it}+1)}{g(N_{it})}
$$

In the specific case of Cournot Competition, $g(N+1)/g(N) = (N +$ $1)^2/N^2$ which decreases less quickly than $(N + 1)/N$. If competition decreases profits at more than rate $1/N$ the true ratio of thresholds will be overestimated. An example is shown in figure 2.3, in which entry thresholds are overestimated with respect to their fixed effect counterparts, but only slightly. Indeed, because competition decreasing profits at rate $1/N$ is such a plausible model, the bias in the threshold ratio will undoubtly be small.

2.4 Panel Data Solution

The panel structure of data can be used to eliminate bias in entry models, as discussed in the context of labor economics by Angrist and Kruger[2]. Unobserved shocks to profitability can be decomposed into:

 $\varepsilon_{it} = m_i$ (market effect) + y_t (year effect) + v_{it}

a component which remains constant over a market's life (m_i) , a component which represents aggregate shocks common to all markets in a year (y_t) while the remaining unobserved elements are grouped into a mean zero shock v_{it} . Estimates are biased to the extent that v_{it} is correlated with demand and number of firms:

$$
v_{it} = \hat{\delta}X_{it} + \hat{\gamma}N_{it} + \hat{\zeta}_{it}
$$

This correlation is likely much smaller than before. Ultimately, the most convincing solution to this problem is to use an instrumental variable strategy. Find a variable z_{it} which is uncorrelated with unobserved profitability ε , but correlated with demand and number of plants, such that $E[\varepsilon z] = 0$. It is then possible to use GMM to estimate an consistent (if not efficient) model of entry.

2.4.1 Computational Details

Fixed effects are commonly introduced into discrete choice models by conditioning techniques such as Chamberlain's fixed effect logit [CITE]. In the case of ordered probit models with groups of 20 observations, conditioning is computationally difficult. Instead, a dummy variable per market is added to the model, estimated using maximum likelihood as any other demand parameter:

$$
\pi_{it} = X_{it}\beta + \sum_{k=1}^{K} 1(k=m)\alpha_m + g(N_{it})
$$

Maximizing the likelihood involves solving for the value of over 3000 parameters, given the number of markets in the data. Fortunately, the linear objective function along with the structure of an ordered probit yields a globally concave likelihood function. This makes this problem computationally feasible since globabally concave function are straighforward to maximize. As well, the gradient of the likelihood is computed analytically, bypassing the computation of a rather large number of numerical derivatives. Finally, the market level fixed effect parameters are îincidentalî in the sense that their values are not of interest, just the effect they have on economically important parameters such sunk costs and the effects of competitors. The termination criteria reflects this,

requiring only that the likelihood to converge $L(\theta_t) - L(\theta_{t-1}) < \varepsilon$ rather than the full vector of parameters: $\|\theta_t - \theta_{t-1}\| < \delta$. The number of iterations required to compute the solution of the model is reduced from 50 to about 5 without changing the value of economically relevant parameters. On a UNIX server, finding the fixed effect maximum likelihood parameters takes approximately a day, but is faster for subsamples of markets.

2.4.2 Random Effects Model

Another method for eliminating bias is think of ε 's as correlated accross time. Assume that ε 's evolve according to the following process:

$$
\varepsilon_{mt} = \rho \varepsilon_{mt-1} + \mu_{mt}
$$

where μ_{mt} is a mean-zero, uncorrelated shock and ρ is the correlation coefficient of unobservables. In the first period, the standard Bresnahan-Reiss model is used, and in subsequent period's. A simulation estimator can be used for this model.

3 Market Definition

An important issue in the decision to enter or exit a market is the competitive environment a firm faces, and in particular the number of competitors in the market which erode its profits. I propose two market definitions for the ready-mix concrete sector: county markets and isolated town markets.

3.1 County Level

The first market definition is the county market: every county in the United States considered a separate market. This approach suffers from several problems. First, counties vary greater in population and hence a county such as Cook could be described as being composed of several different submarkets while smaller rural counties such as Brown (IL) are components of a larger regional market. Second, most counties in the United States were constituted long before settlement, and hence they are often not centered on a central place (except for particularly rural counties where the county seat is the only agglomeration of any importance). Third, some counties in the Western States tend to be very large (in particular Arizona has only 12 counties while Illinois has almost one hundred). With these proviso's in mind, counties still form a useful benchmark for market definition since they are consistently coded in Census dataset and have not changed geographically in the last 50 years (with a handful of notable exceptions).

We have also constructed a table of neighboring counties using 3 different but nested measures:

- Adjacent Counties: Counties that border this county (up to 12).
- Within 20 miles: Counties whose nearest point is within 20 miles of this county (up to 17).
- Within 30 miles: Counties whose nearest point is within 30 miles of this county (up to 24).

These are constructed to try to control for the influence of out of county plants. In particular, there may be problems with counties located near very large population centers having unusually large demand.

3.2 Place Selection

A more natural approach to market definition is to focus on isolated towns. This method was used by Bresnahan and Reiss[?] in their pioneering study of entry barriers in isolated retail activities. If anything concrete offers a more compelling basis for market segmentation since transportation costs are high and transit time is restricted to less than an hour because of the hardening of the concrete mixture. This creates hard market segmentation, which is alleviated only by adding of certain chemical compounds (known as admixtures) which can moderately increase setting times by about 20%. For the purposes of this study, two markets are distinct if they are:

- Over 4000 inhabitants in 1990.
- More than 20 miles from a place with more than 4000 inhabitants in 1990.
- More than 30 miles from a place with more than 4000 inhabitants in 1990 if the place has an interstate running through it.

(Towns within 1 mile of each other are considered as a single unit.)

An example of an isolated town is Tuba City, Arizona whose isolation is illustrated by figure 4. While these screens are somewhat ad hoc, the conceptual basis for them is far cleaner than in (Bresnahan and Reiss 1990): concrete deteriorates fairly quickly as it is transported

	Minimal Size of Town to be considered	
Minimal Distance to Closest Town \parallel 4000 Population \mid 2000 Population		
20 miles	375	404
30 miles	160	167
40 miles		

Table 1: Number of Isolated Markets with Different Screens

away from the plant. An arbitrary sample of adjacent towns were used in http://www.mapquest.com/ to test the link between driving time and distance. In the vast majority of cases, a distance of 20 miles corresponds to driving times of over 45 minutes. These screens produce the following number of towns shown in table 1.

3.3 Zip Level

However, information about isolated towns does not lead us directly to a market. While occupations such as dentists[5] are principally concentrated in towns, ready-mix concrete operators frequently tend to locate on a towns outskirts, frequently outside the municipality proper, which is the geography coded in the CMF. To rectify this problem, plants and construction establishments are selected by Zip Code. The basic criterion for inclusion of a zip code in the town/market is that a zip code must be within 5 miles of town limits. The case of Tuba City for zip codes is shown in figure 3.

There are also a few issues with using Zip codes to define a market. First, Zip codes are frequently of an irregular shape due to their primary purpose: efficient delivery of mail, following roads and going around rivers. Second, Zip codes occasionally change due to the US postal service rearranging its postal delivery scheme. On the other hand, Zip codes are typically centered on these isolated towns, and hence form a better fit to the idea of market than might be expected.

3.4 Static Entry Estimates

Static Entry Models have been estimated both for county markets and zip markets. Since the coefficients for these two models are statistically indistinguishable, we will focus presenting county market evidence where there are fewer issues with shifting market definitions.

Estimating a fixed effect Bresnahan-Reiss model no simple undertaking. Fixed effect models for discrete data tend to rely on conditioning on the total number of events choosen in a group's lifetime. However, for an ordered probit model, this implies creating groups of the times the number of firms ranging from 0 to 20 was choosen. The total number of groups quickly grows, making it difficult to estimate the probability that this number of groups occurs given the parameter vector. Instead, coefficients for each market are estimated, leading to a model with over 2000 parameters.

Results in table 2 for Bresnahan-Reiss Entry model and table 3 for Sunk-Cost Bresnahan-Reiss Estimator. Note that the coefficient on de-

mand is more than halved and the coefficient on number of competitors becomes twice as negative when fixed effects are added to the model.

If we care about the reaction of the number of establishments to changes in demand, such as recessions. In this application, we really care about distinguishing Öxed market characteristics such as population from employment in the construction sector. As is shown in figure 2, adding fixed effects dramatically lowers the response of a market to demand shocks.

References

- [1] Aguirregabiria, Victor and Pedro Mira (2004) "Sequential Estimation of Dynamic Discrete Games" working paper Boston University.
- [2] Angrist, Joshua D. and Alan B. Krueger (1999) "Empirical Strategies in Labor Economics." Handbook of Labor Economics, vol. 3A, edited by Orley Ashenfelter and David Card, Amsterdam:Elsevier.
- [3] Bajari, Patrick and Lanier Benkard, Jon Levin "Estimating Dynamic Models of Imperfect Competition" working paper Duke University.
- [4] Bresnahan, T. F. and P. C. Reiss (1990). "Entry in Monopoly Markets." Review of Economic Studies 57(4): 531-53.
- [5] Bresnahan, T. F. and P. C. Reiss (1994). "Measuring the Importance of Sunk Costs." Annales díEconomie et de Statistique 0(34): 181-217.
- [6] Collard-Wexler, Allan (2004) "Constructing Longitudinal Data on the Ready-Mix Concrete Sector and Concrete Demand", working paper Northwestern University.
- [7] Jarmin, R. S. and J. Miranda (2002). The Longitudinal Business Database, US Census, Economic Studies.
- [8] Kosmatka, S. H., B. Kerkhoff, et al. (2002). Design and control of concrete mixtures. Skokie, Ill., Portland Cement Association.

4 Appendix

Figure 2: Entry and Exit Thresholds for Standard Model and Fixed Effect Model.

Table 2: Bresnahan-Reiss Estimates with and without county fixed effects

Table 3: Standard and Fixed Effect Sunk Cost Bresnahan-Reiss Estimates for County Markets

Figure 3: Zip Codes around Tuba City.

Figure 4: Isolated Place: Tuba City, Arizona.

	Parameter	Standard Error
Constant	0.56	(0.09)
Inhabitants (in millions)	7.57	(1.60)
Vacant Housing (in millions)	-1.49	(2.60)
Interstate within 5 miles	-0.24	(0.11)
1 competitor	-1.10	(0.07)
2 competitor	-0.79	(0.07)
3 competitor	-0.65	(0.09)
4 competitor	-0.27	(0.09)
More than 4 competitors	-0.25	(0.07)
Number of observations	449	
Wald $\text{chi2}(3)$	25.16	
Log Likelihood	-653	

Table 4: Bresnahan-Reiss Model for Zip Markets

Figure 5: Zip and County Markets give effects of competition. Bars represent 95% confidence intervals on parameter estimates.