Demand Fluctuations in the Ready-Mix Concrete Industry

Allan Collard-Wexler*
NYU Stern

August 13, 2010

*The work in this paper is drawn from chapter 2 of my Ph.D. dissertation at Northwestern University under the supervision of Mike Whinston, Rob Porter, Shane Greenstein and Aviv Nevo. I would like to thank the anonymous referees for comments which greatly improved the paper and to Lynn Riggs, Ariel Pakes, John Asker, Mike Mazzeo and Ambarish Chandra for helpful conversations. The Fonds Québécois de la Recherche sur la Société et la Culture (FQRSC) and the Center for the Study of Industrial Organization at Northwestern University (CSIO) provided financial support. I would like to thank seminar participants at many institutions for comments. The research in this paper was conducted while I was a Special Sworn Status researcher of the U.S. Census Bureau at the Chicago Census Research Data Center. Research results and conclusions expressed are those of the author and do not necessarily reflect the views of the Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Support for this research at the Chicago and New York RDC from the NSF (awards no. SES-0004335 and ITR-0427889) is also gratefully acknowledged.

Abstract

Census Data on over 15 000 ready-mix concrete plants, from 1976 to 1999, is used to investigate the role of demand shocks. A model of investment and entry in oligopolistic markets is estimated using a Conditional Choice Probability Indirect Inference Algorithm implemented via a Stochastic Algorithm. Estimates from this model are used to simulate the effect of eliminating short-term (i.e. 5 year) demand changes at the county level. Adjusting government expenditures in a way that reduces volatility of demand has a "market expansion" effect. The model predicts a 47% increase in the number of firms in the industry and an increase in investment of 52% from $0.95 billion dollars per year to $1.6 billion per year. Since bigger markets have both more plants and larger plants, the demand smoothing fiscal policy would increase the share of large plants in the industry by 41%. However, I find no effect of demand smoothing on either the entry or exit rate or adjustments to plant size.

1 Introduction

Many industries face considerable uncertainty about future demand for their products. How do these shocks affect the organization of production and market structure?

I study the effect of demand shocks in the ready-mix concrete industry. Concrete is a geographically segmented industry since wet concrete cannot travel much more than an hour before hardening. Thus the industry is composed of local oligopolies and demand shocks cannot be deflected by reallocating production to other markets. The ready-mix concrete industry witnesses large changes in demand from the construction sector from year to year which are of great concern to ready-mix producers, as the size of the construction industry at the county level changes on average by 30% per year. Moreover, about half of concrete is used by state and local governments to construct roads and buildings. These government outlays are volatile due to year to year changes in tax revenues.

I look at government intervention in the ready-mix concrete market that would smooth out short-term fluctuations in demand at the county level. Specifically, the counterfactual mimics the effect of government sequencing its contracts in such a way to fix demand for a five year period. After 5 years are up, demand changes in the same way as it would have absent the demand smoothing policy.

I use longitudinal data provided by the Center for Economic Studies at the

U.S. Census Bureau, on the life histories of over 15 000 ready-mix concrete plants in United States from 1976 to 1999 to simulate the effects of this policy. Because the ready-mix concrete industry is an oligopoly, I estimate a model of entry and discrete investment in concentrated markets using an Indirect Inference Conditional Choice Probability Algorithm. Moreover, I use the Stochastic Algorithm of Pakes and McGuire (2001) to deal with very large state spaces. This allows for considerable plant heterogeneity.[1]

Plant size is directly related to market size in the ready-mix concrete sector: bigger markets have both more plants and larger plants. Estimates of the model show that construction employment has strong positive effects on profits but disproportionately so on large plants. Competitors in the market - in particular the first competitor - substantially reduce plant profits and there are large sunk costs both for entering the market and for increasing or shrinking the size of a plant, and these costs are greater for large plants than for small plants.

The counterfactual policy of smoothing out 5 year changes in demand has substantial effects on the industry. Smoothing demand would increase the number of plants in the industry by 47%. Investment would increase by 52% from $0.95 billion to $1.6 billion per year, while producer surplus for incumbents would increase by 29%. Moreover, the share of large plants (with

[1]Previous versions of the paper used an algorithm analogous to Aguiregabiria and Mira (2007) where the choice probabilities were updated to match those given by a computed equilibrium of the game given the estimated parameter vector. This technique leads to similar estimates and counterfactual results as those presented in the paper and are available by request.

more than 17 employees) would rise by 41%.

Reducing demand volatility has a "market expansion" effect. I find that that a 1% increase in market size (as measured by construction employment) is associated with a 0.69% increase in the number of ready-mix concrete plants. The concavity of the effect of construction employment implies that a reduction in the volatility of demand increases the number of firms in the industry. As the "market expansion effect" is similar to an increase in market size, the size distribution of the industry will shift towards large plants, since the profits of these large plants increases more quickly with market size than for small plants. The market expansion effect reduces the number of monopoly markets from 71% of counties to 42% of counties, while the number of counties in which there is no ready-mix concrete plant remains fixed at 2%. I find that prices are 4% lower in competitive markets than monopoly markets, so consumers would spend 64 million dollars per year less on ready-mix concrete. Finally, producers would benefit from demand smoothing, as the net present value of producer surplus would rise from 3.3 to 4.0 billion dollars.

Demand shocks have little effect on entry and exit rates or the rate at which plants change their size. High sunk costs of entry and adjustment make firms unlikely to respond to temporary changes in demand. As demand volatility shrinks, firms get a more precise forecast of future demand and thus changes in demand become more relevant. The direct effects of a smoother demand process is exactly offset by firms becoming more sensitive to demand, removing any effect of demand smoothing on turnover and the rate at which plants adjust their size.

In section 2, I discuss the source of sunk costs for ready-mix plants, and the role of spatial differentiation in the industry. Section 3 describes how I construct the data. In section 4, I present a dynamic model of competition, and I describe estimation in section 5 and results in section 6. Finally, in section 7 I discuss the effect of policies that would eliminate some of the volatility of demand.

2 The Ready-Mix Concrete Industry

Concrete is a mixture of three basic ingredients: sand, gravel (crushed stone) and cement, as well as chemical compounds known as admixtures. Combining this mixture with water causes the cement to undergo an exothermic chemical reaction called hydration, turning cement into a hard paste that binds the sand and gravel together. I focus on ready-mix concrete: concrete which is mixed with water at a plant and transported directly to a construction site. Ready-Mix is a perishable product that needs to be delivered within an hour and a half before it becomes too stiff to be workable.[2] Concrete is also very cheap for its weight. One producer describes the economics of transportation costs in the ready-mix industry as follows:

[2] "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" p.96 in Kosmatka, Kerkhoff, and Panarese (2002).

> "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. That means that concrete is cheaper than shit. [3]"

A ready-mix truck typically drives 20 minutes to deliver a load. [4] Thus, concrete's most salient feature from an economic perspective is that markets are geographically segmented. Figure 1 shows the dispersion of ready-mix producers in the Midwest, with an handful of incumbents in each area. In my empirical work I treat each county as a separate market, one that evolves independently from the rest of the industry.

Table 1 shows that the vast majority of counties in the United States have fewer than 6 ready-mix plants, reflecting a locally oligopolistic market structure. At the same time, because even the most isolated rural areas has demand for ready-mix concrete, most counties are served by at least one ready-mix producer.

A market with more than 3 firms appears to yield fairly competitive outcomes. Figure 2 shows the median price in markets with 1 to 7 firms in

[3]Telephone interview, January 2005.

[4]The driving time of twenty minutes is based on a dozen interviews conducted with Illinois ready-mix concrete producers. Thanks to Dick Plimpton at the Illinois Ready-Mix Concrete Association for providing IRMCA's membership directory.

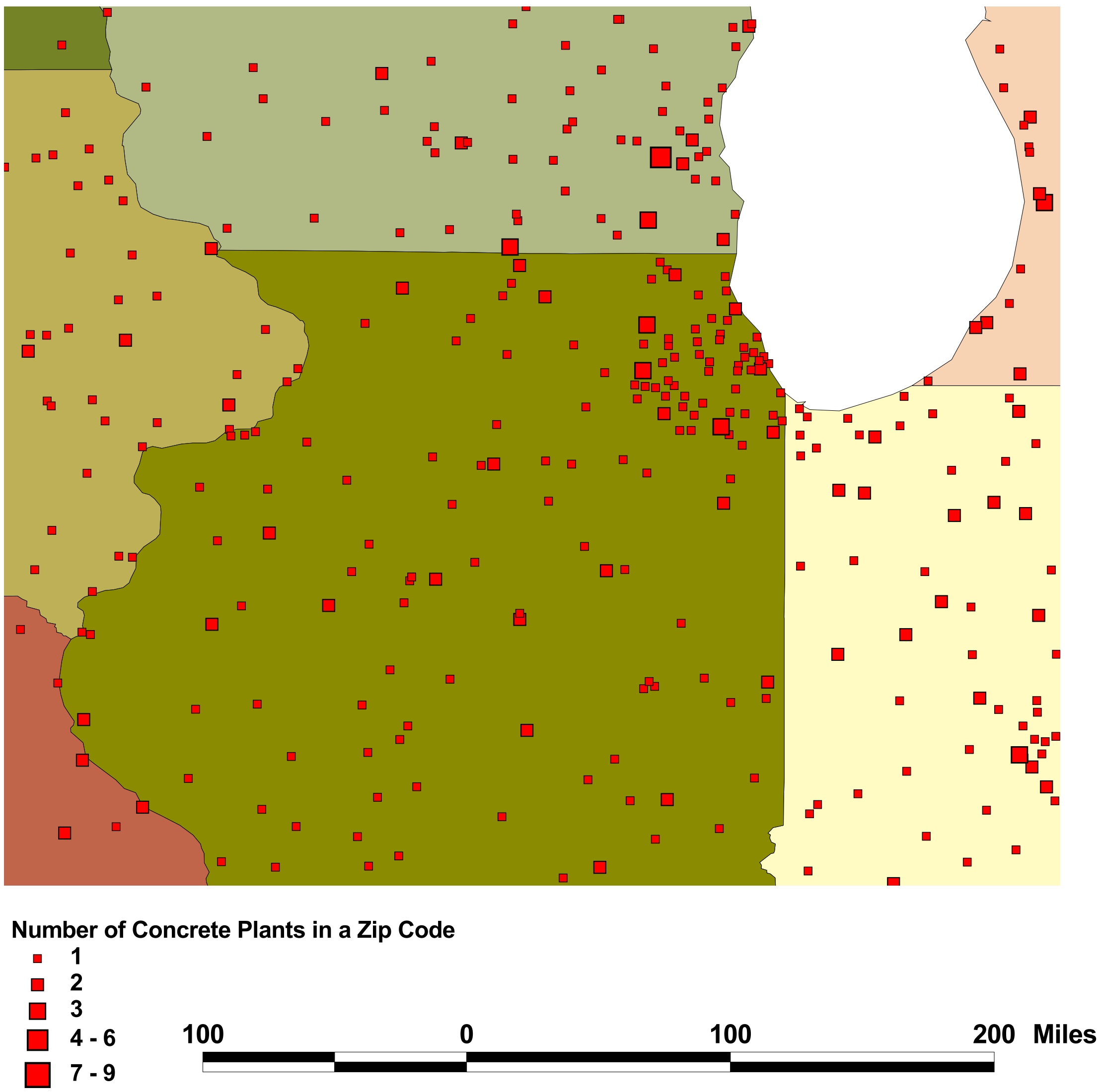


Figure 1: Dispersion of Ready-Mix Plant Locations in the Midwest.
Source: Zip Business Patterns publicly available dataset at
http://www.census.gov/epcd/www/zbp_base.html.

Number of Concrete Plants	Number of Counties/Years	Percent
0	22,502	30%
1	23,276	31%
2	12,688	17%
3	6,373	9%
4	3,256	4%
5	1,966	3%
6	1,172	2%
More than 6	3,205	4%
Total	74,438	

Table 1: Most counties in the United States are served by less than 6 ready-mix concrete plants.

the county.[5] Note that the first 3 competitors have a noticeable impact on prices, but beyond this there is little decline in prices. Price is constructed similarly to Syverson (2004a), i.e. price is sales of concrete divided by tons of concrete sold in the material trailer of the Census of Manufacturing.

Ready-Mix concrete is essentially a homogeneous good. While it is possible to produce several hundred types of Ready-Mix concrete, these mixtures basically use the same ingredients and machinery. Because of aggressive antitrust policy on the part of the Department of Justice, the typical ready-mix producer is a single plant operator. Indeed, Syverson (2004a) reports that 3749 firms controlled the 5319 ready-mixed plants operating in 1987. Thus I will assume that each firm owns a single ready-mix concrete plant, making plant and firm interchangeable.

[5] These prices have been constructed using sales of concrete divided by volume of concrete, following Syverson (2004a) procedure which removes hot and cold deck imputes by dropping all price pairs which are exactly the same. The appendix discusses more details on the construction of price statistics.

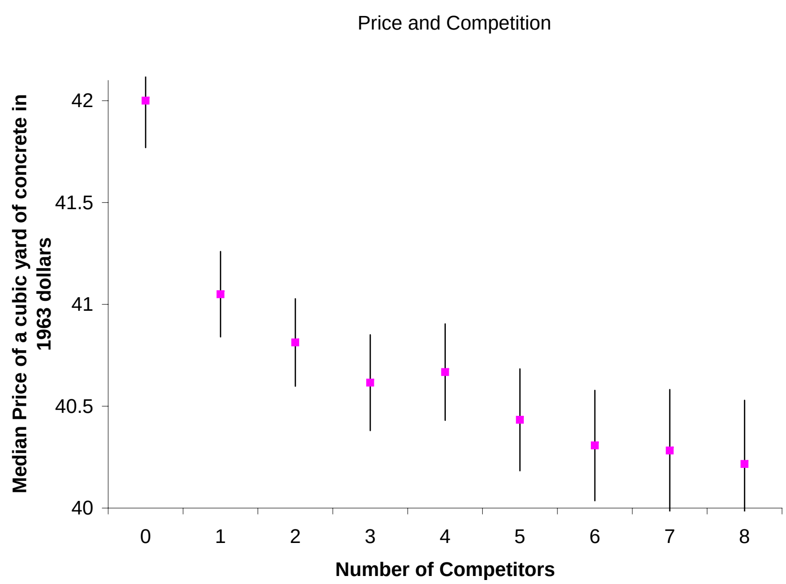


Figure 2: Price declines with the addition of the first competitors, and little afterwards.
Note: Bars represent 95% confidence interval on median price.

Concrete is consumed by the construction sector. The bulk of concrete purchases are made by the construction sector, to build apartments, houses, roads and sidewalks. I use employment in the construction sector as my demand measure. Demand for ready-mix concrete is inelastic since it is a small part of construction costs. Concrete costs do not exceed 10% of material costs for any sector in construction. So it is unlikely that the ready-mix market substantially affects the volume of construction activity. Government purchases about half of U.S. concrete, primarily for road construction. According to the Kosmatka, Kerkhoff, and Panarese (2002) p.9, Government accounts for 48% of cement consumption, with road construction alone responsible for 32% of total consumption. Fluctuations in Government purchases of concrete are mainly due to the discretionary nature of highway spending in state and federal budgets. These purchases are a major source of uncertainty for ready-mix producers.[6]

The autocorrelation of log county construction employment is 85% for 1 year, 74% for 2 years, 65% for 5 years and 21% for 20 years. This is a very low autocorrelation for construction activity, signaling an enormous amount of year to year change in demand. However, the demand process has more long term correlation than an AR process would predict (note that an AR(1) process would predict a 4% 20 year autocorrelation given a 85% 1 year autocorrelation). To capture long-run differences in market size, I split up markets

[6]Conversation with Edward Sullivan, chief economist at the Portland Cement Association, May 2005.

	Change in Log Construction Employment		
	1 Year Change	5 Year Change	10 Year Change
Change in Log Construction Employment			
In adjacent counties	1.1%	2.1%	2.8%
In counties within 10 miles	1.1%	2.4%	3.2%
In counties within 20 miles	0.6%	2.4%	3.4%

Table 2: Spatial Correlation of Changes in Log Construction Employment.

into categories based on rounding the average number of firms in a market to the nearest integer. I then estimate the demand process separately for each market size category. [7]

Table 2 shows regressions of changes in log county construction employment on changes in nearby counties. These regressions indicate very little spatial autocorrelation in demand. For instance, only 2.1% of the variation in log construction employment in a county is accounted for by changes in log construction employment in counties that border it. Moreover, any aggregate component of construction employment would show up as spatial autocorrelation of changes in construction activity. At the county level we can think of demand evolving autonomously, and I consider policies that would interfere with county level demand patterns instead of state or national patterns.

Opening a concrete plant is an expensive investment. In interviews, managers of ready-mix plants estimate the cost of a new plant at between 3 and

[7] In Appendix A I discuss alternate proxies for persistent differences between markets. The number of firms in the market is an endogenous variables- thus problematic - I find similar first-stage entry and exit regressions when I categorize markets using non-endogenous variables such as the lagged number of firms in the market, or market fixed-effects, indicating that in practice endogeneity is a minor problem. However, using the lagged number of firms as a categorization variable makes it possible for a market to switch from one category to another, adding a layer of complexity in the state space that I wish to avoid.

4 million dollars, and continuing plants in 1997 had on average 2 million dollars in capital assets. There are few expenses involved in shutting down a ready-mix plant. Trucks can be sold on a competitive used vehicle market, and land can be sold for other uses. 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 outweighs the benefits. I provide evidence of sunk costs in the ready-mix industry at the plant level, including factors difficult to quantify, such as long term relationships with clients and creditors. These intangible assets may account for a large fraction of sunk costs. [8]

Ready-mix concrete has been studied extensively by Syverson (2004a), who provides evidence of productivity dispersion across plants. This productivity dispersion is evidence of large differences between plants which are not eliminated by competitive pressures. I provide an explanation for why the competitive adjustment process is not instantaneous.

3 Data

Data on Ready-Mix Concrete plants are taken from data provided by the Center for Economics Studies at the United States Census Bureau. The primary data source is the Longitudinal Business Database (henceforth LBD)

[8]For instance, ready-mix operators sell about half of their production with a six month grace period for repayment. Accounts receivable have a value equivalent to half of a plant's physical capital assets. It will be more difficult to collect these accounts if the firm cannot punish non-payment by cutting off future deliveries of concrete.

compiled from data used by the Internal Revenue Service to maintain business tax records. The LBD covers all private employers on a yearly basis from 1976 to 1999 and has information about employment, salary along along with sectoral coding and certain types of business organization data such as firm identification.

Production of ready-mix concrete for delivery predominantly takes place at establishments in the ready-mix sector. Hence, establishments in the ready-mix sector are chosen, corresponding to either NAICS (North American Industrial Classification) code 327300 or SIC (Standard Industrial Classification) code 3273.

To construct longitudinal linkages, I adapt the Longitudinal Business Database Number (henceforth LBDNUM), as developed by Jarmin and Miranda (2002). This identifier is constructed from Census ID, employer ID and name and address matches of all plant in the LBD. To identify plant entry and exit, I use Jarmin and Miranda (2002)'s plant birth and death measures. Jarmin and Miranda identify entry and exit based on the presence of a plant in the I.R.S.'s tax records. They take special care to flag cases where plants simply change owners or name by matching the address of plants across time.[9] Each year, about 40 plants (or about 1.6% of plants) are temporarily shut down. I do not treat temporary shutdown as exit, since the cost of reactivating a plant is far smaller than building one from scratch.

[9] If a plant changes ownership, I do not treat this as an exit event since the cost of changing the management at a plant should be much lower than the cost of building a plant from scratch.

I complement the LBD with data from the Census of Manufacturing (henceforth CMF) and Annual Survey of Manufacturers (henceforth ASM) which contains more detailed information on plants such as inputs, outputs and assets. Unfortunately, the ASM is only sent to about a third of plants in the ready-mix concrete sector, while the CMF is available only every five years and excludes all plants with less than 5 employees (about a quarter of concrete plants). Since the CMF and ASM have serious issues with missing data, it makes it difficult to use them alone for longitudinal market level studies. This is not true of the LBD which has the entire population of plants. Finally, construction data is obtained by selecting all establishments from the LBD in the construction sector (SIC 15-16-17) and aggregating them to the county level.

3.1 Panel

Plants occasionally switch in and out of the ready-mix concrete sector. I select all plants that have belonged to the ready-mix sector at some point in their lives, but throw out plants that switch into the concrete sector for a small fraction of their lives, since these transient concrete plants are typically miscoded plants manufacturing other products such as cement of concrete pipe.[10] As a product, ready-mix concrete represents 95% of shipments from my selection of concrete plants. Moreover, when I collect all plants that

[10] Specifically I toss plants that produce concrete less than 50% of the time from my sample.

produce ready-mix concrete, based on their response to the product trailer of the Census of Manufacturing (which collects detailed information on the output of plants), I find that 94% percent of ready-mix concrete is produced by plants in my sample.

Table 3 shows that over the sample period there are about 350 plant births and 350 plant deaths each year compared to 5000 continuers. Turnover rates and the total number of plants in the industry are fairly stable over the last 30 years. The average ready-mix concrete plant employs 26 workers and sold about 3.4 million dollars of concrete in 1997, split evenly between material costs and value added. However, these averages mask substantial differences between plants. Most notably, the distribution of plant size is heavily skewed, with few large plants and many small ones, indicated by the fact that more than 5% of plants have 1 employee, while less than 5% of plants have more than 82 employees. Continuing firms are twice as large as either entrants(births) or exitors(deaths), measured by capitalization, salaries or shipments. Plant size (as measured by employment) is the most important plant state since bigger plants ship far more concrete and are much less likely to exit. Salaries have a correlation of 92% with total shipments (versus 43% for capital) and are very persistent, with an autocorrelation of 91% (versus only 74% for capital). For this reason the number of employees is used as measure of plant size.[11]

[11] I use employment instead of capital stock, since employment is measured for all plants in the data (since it is derived from IRS tax returns in the LBD), while capital is available for all plants in a market for only a small number of markets (as is discussed in Collard-

I aggregate plant data by county to form market level data. Since counties in the United States vary greatly in size, I have taken care to exclude counties in states such as Arizona which have unusually spacious counties, and a small number of heavily populated urban counties.[12] Table 4 presents summary statistics of the market level data. Note that on average there are 1.86 plants per market. Moreover the is a wide range of construction employment, going from 11 employees (5^{th} percentile) to 6800 employees (95^{th} percentile), 500 times greater. Yet the range of the surface of county in square miles is 210 to 3200 a bit more than a 10 times difference.

Wexler (2009) which use multiple imputation to fill in missing capital stock). In practice, given the coarseness the size of my employment bins, classifying a firm based on capital or employment does not matter very much.

[12]Specifically, counties with more than 20 plants. The County Business Patterns reports that in 2007 there are 20 of these counties.

Year	Birth	Continuer	Death
1976	501	4,737	N.A.
1977	557	4,791	410
1978	327	5,043	445
1979	392	5,093	333
1980	271	5,140	387
1981	313	5,069	360
1982	313	4,875	423
1983	273	4,991	315
1984	328	4,972	295
1985	309	4,988	339
1986	300	5,003	305
1987	390	4,898	404
1988	270	5,016	269
1989	248	4,275	448
1990	194	4,103	304
1991	220	3,882	291
1992	214	4,643	348
1993	133	3,668	270
1994	163	3,952	232
1995	196	3,840	243
1996	195	3,734	230
1997	338	4,768	274
1998	239	4,949	267
1999	320	4,961	234

Table 3: The number of Births, Deaths and Continuers is fairly stable over the last 25 years

	Observations	Mean	Standard Deviation	5th Percentile	95th Percentile
County Total Value of Shipment (in 000's)	24677	3181	12010	0	14000
County Value Added (in 000's)	24677	1408	5289	0	6500
County Total Assets Ending (in 000's)	24677	1090	14134	0	4700
County Concrete Plants	74435	1.86	3.24	0	6
County Employment	74435	27.24	79.03	0	110
County Payroll (in 000's)	74435	4238	74396	0	3600
County 0-5 Employee Plants	74435	0.52	1.07	0	2
County 5-20 Employee Plants	74435	0.78	1.34	0	3
County more than 20 Employee Plants	74435	0.86	1.49	0	3
County Plants Less than 5 years old	74435	0.17	0.76	0	1
County Plants 5-10 Years Old	74435	0.54	1.47	0	2
County Plants over 10 Years Old	74435	1.35	2.07	0	4
County Area	72269	1147	3891	210	3200
Employment in Construction	69911	1495	5390	11	6800
Payroll in Construction (in 000's)	69911	37135	163546	110	160000

Table 4: Summary Stats for County Aggregate Data

4 Model

I use the theoretical framework for dynamic oligopoly developed by Ericson and Pakes (1995)- with some small variations - to analyze entry, exit and investment decisions in the ready-mix concrete industry. In each market there are $i = 1, \cdots, N$ firms in the market, which are either potential entrants or incumbents. A firm i can be described by a firm specific state $s_i^t \in S_i$. Firms also react to market-level demand, M^t, and thus the market level state s^t is the composition of the states for each firm and the aggregate state M^t:

$$s^t = \{s_1^t, s_2^t, \cdots, s_N^t, M^t\}$$

I will distinguish between two component of the state s_i^t, x_i^t which is common knowledge to all firms in the market and ε_i^t which is an i.i.d. private information component. If instead ε_i^t was serially correlated, then a firm might find it optimal to condition it's strategy on past actions taken by other firms in the market. This would seriously increase the size of the state space [13] Denote by $x^t = \{x_1^t, x_2^t, \cdots, x_N^t, M^t\}$ and $\varepsilon^t = \{\varepsilon_1^t, \varepsilon_2^t, \cdots, \varepsilon_N^t\}$ the market level common knowledge and private information state respectively.

In each period t, potential entrants choose of whether to enter a market, and incumbents can choose to exit the market. Moreover, conditional on being in the market, firms pick their common knowledge state x_i^t in the next

[13] In my empirical application, there are 1.4 million states. If a firm also conditioned on the history of the market for even a single year, then the state space would be over 1.9 billion billions.

period. Thus the firm's action a_i^t is the choice of being out of the market, i.e. $x_i^{t+1} = \emptyset$, or their state tomorrow x_i^{t+1} conditional on choosing to have a plant in the market. Demand evolves following a first-order Markov Process transition probabilities given by $D(M^{t+1}|M^t)$.

I assume the private information component ε_i^t enters into the profit function as an additive shock to the value of each action a_i^t. Thus payoffs are given by:

$$r(x^{t+1}) + \tau(x_i^{t+1} = a_i^t, x_i^t) + \varepsilon_{ia}^t \tag{1}$$

where $r(\cdot)$ denote the rewards from operating in the market, and $\tau(\cdot)$ are transition costs, i.e. the costs of moving from one state to another. The results section of the paper are primarily concerned with estimating these reward and transition functions.

The timing of the game is thus:

1. Firms privately observe ε_i^t and publicly observe x^t.

2. Firms simultaneously choose actions a_i^t.

3. Demand M^t evolves to it's new level M^{t+1}. Firm level states evolve to x_i^{t+1}.

4. Payoffs $r(x^{t+1}) + \tau(a_i^t, x_i^t) + \varepsilon_{ia}^t$ are realized.

So I can define the firm's ex-ante (i.e. before observing ε_i^t) value as:

$$V(x^t) = \mathbb{E}_{\varepsilon_i^t}\left(\max_{a_i^t} \mathbb{E}_{x_{-i}^{t+1}}\left[r(x_{-i}^{t+1}, x_i^{t+1}) + \tau(x_i^{t+1} = a_i^t, x_i^t) + \varepsilon_{ia}^t + \beta V(x^{t+1})\right]\right) \tag{2}$$

and actions are chosen so as to maximize the net present value of rewards:

$$a_i^t = \text{argmax}_{a_i^t} \mathbb{E}_{x_{-i}^{t+1}}\left[r(x_{-i}^{t+1}, x_i^{t+1}) + \tau(x_i^{t+1} = a_i^t, x_i^t) + \varepsilon_{ia}^t + \beta V(x^{t+1})\right] \tag{3}$$

A symmetric Nash equilibrium in pure strategies is a set of policies $a^*(x^t, \varepsilon_i^t)$ such that a unilateral one shot deviation to strategy $\tilde{a}_i(x^t, \varepsilon_i^t)$ would not lead to a higher net present value of rewards, conditional on all other players using strategies $a_{-i}^*(\cdot)$. Notice that strategies of other players are simply embedded into the value function $V(\cdot)$ and the expectations over x_{-i}^{t+1}. If ε_{ia}^t is an additive, action specific shock and if it has full support, there will exist pure strategy equilibria to this game. [14]

4.1 Stochastic Algorithm

To compute the strategies associated with a Nash Equilibrium of the dynamic game, I adapt the stochastic algorithm of Pakes and McGuire (2001) to the discrete action setup used in this paper since the state space has up to 1.4

[14] Proposition 2 in Doraszelski and Satterthwaite (2010) describes conditions under which the Ericson and Pakes (1995) model has a pure strategy equilibrium, essentially pointing out that exit and entry costs need to have full support shocks to ensure the existence of pure strategy equilibrium. The game I describe has full support shocks to the value of entering and exiting, as well as to the value of taking any action.

million states.[15]

An important assumption that I will use is the fact that the private information state $\epsilon^t_{a_i}$ is a i.i.d. logit random variable, an assumption that I will defer discussing until to the next section.

First, some notation. To work out the firm's strategies, I will compute the *ex-ante choice specific value function* $W(a_i, x)$, i.e. the net present value of payoffs conditional on taking action a_i (before I observe ϵ^t_i) defined as[16]:

$$\begin{aligned} W(a_i, x) &= \mathbb{E}_{x'|a_i}\left[r(x') + \tau(a_i, x_i) + \beta V(x')\right] \\ &= \mathbb{E}_{x'|a_i}\left[r(x') + \tau(a_i, x_i) + \beta \mathbb{E}_{\epsilon'}\left(\max_{a'_i} W(a'_i, x') + \epsilon'_{a_i}\right)\right] \end{aligned} \tag{4}$$

Given the choice-specific value function, it is easy to reckon the firm's conditional choice probability (henceforth CCP) $\Psi[a_i|x]$, i.e. the probability that a firm will play action a_i in observable state x - before observing ϵ_i - using the logit formula:

$$\Psi[a_i|x] = \frac{\exp\left(W(a_i, x)\right)}{\sum_{j \in A_i} \exp\left(W(j, x)\right)} \tag{5}$$

Next I define the *hit counter*, denoted $h(a_i, x)$, as the number of times the location $l = (a_i, x)$ has been visited by my algorithm. The hit counter

[15]There are 10 firms, 7 possible states per firm and in the most complex model 50 demand states. I reduce the size of the state from $10^7 \times 50$ to 1.4 million by using the assumption of exchangeability described by Gowrisankaran (1999).

[16]Note that V and W are linked together by:

$$V(x) = \mathbb{E}_{\epsilon_{a_i}}\left(\max_{a_i} W(a_i, x) + \epsilon_{a_i}\right)$$

is important since it will allow me to keep track of the precision of the computation of $W(a_i, x)$ and $\Psi[a_i|x]$ using the Discrete Action Stochastic Algorithm (henceforth DASA).

Algorithm Discrete Action Stochastic Algorithm (**DASA)**

1. Start in a location $l^0 = \{a_i^0, x^0\}$.

2. Draw an action profile for other players $a_{-i} \sim \prod_{k \neq i} \Psi[a_k|x]$. Given the action profile $a = \{a_i, a_{-i}\}$ draw a state in the next period x':

$$x'|a \sim \hat{D}[M'|M] \prod_i \iota(x_i'|a_i, x_i) \tag{6}$$

   where $\iota(x_i'|a_i, x_i)$ is the *updating* function, which updates the firm's state based on a firm's action and the firms largest size in the past. [17]

3. Increment the hit counter (how often you have visited the state-action pair): $h'(a_i, x) = h(a_i, x) + 1$.

[17] Later in the paper I will make the firm's previous state relevant to the transition cost. Specifically, if the firm's state is $x_i = \{x_i^C, x_i^P\}$, i.e. the current size x_i^C and the largest size in the past x_i^P, then the updating function $\iota(x_i'|a_i, x_i)$ is:

$$x_i' = \begin{cases} \{a_i, a_i\} \text{ if } a_i \geq x_i^P \\ \{a_i, x_i^P\} \text{ if } a_i < x_i^P \end{cases}$$

4. Compute the value R of the action as:

$$
\begin{aligned}
R = & r(a_i, x) - \tau(a_i, x_i) \\
& + \beta \sum_{j \in A_i} W(j, x')\Psi[j|x'] + \beta E(\varepsilon|x', P)
\end{aligned}
\tag{7}
$$

where $E(\varepsilon|x', P) = \left(\gamma - \sum_{j \in A} \ln(\Psi[j|x'])\Psi[j|x']\right)$ (where γ is Euler's Constant).

5. Update the W-function:

$$
W'(a_i, x) = \alpha[a_i, x]R + (1 - \alpha[a_i, x])\, W(a_i, x) \tag{8}
$$

where $\alpha = \frac{1}{h(a_i, x)}$.[18]

6. Update the Policy Function Ψ for state x:

$$
\Psi'[a_i|x] = \frac{\exp\left(W(a_i, x)\right)}{\sum_{j \in A_i} \exp\left(W(j, x)\right)} \tag{9}
$$

for all actions $a_i \in A$.

[18] The main problem with the stochastic algorithm is: 1- making sure the entire state space is searched, 2- ensure fast learning about the W function at the start of the algorithm and 3- making sure that the convergence properties of the algorithm are satisfied. First, I initialize the starting W using fairly high values so that the algorithm visits all states before lowering the estimate of W. Second, at the start of the algorithm I use $\alpha = 1/\sqrt{h(a_i, x)}$ to ensure that initially inaccurate W's get updated quickly. As well, I reset the hit counter after 20 million iterations to ensure that the first rounds of updates are down-weighted. Third, in the final stage of the algorithm I switch to the $\alpha = 1/h(a_i, x)$ update rule which satisfies the convergence properties in Pakes and McGuire (2001) and more broadly and related to the properties of Stochastic approximation algorithms decribed in Powell (2007) on page 216.

7. Draw a new action $a'_i \sim \Psi[\cdot|x']$.

8. Update current location to $l' = \{a'_i, x'\}$.

9. The stopping rule for this algorithm is based on Fershtman and Pakes (2004) which compares the W-function to a simulated average based on rewards from steps 2 and 4 for states that are recurrent. If the W-function is exact, then the squared difference between between these two objects (weighted by the ergodic distribution) can be accounted for by simulation error. The stopping rule is presented in appendix B.

5 Conditional Choice Probability (CCP) Estimation

Applying the Ericson and Pakes (1995) framework to data has proven difficult due to the complexity of computing a solution to the dynamic game. Even with the DASA presented in the previous section it takes more than an hour to compute a solution. For single agent problems, Hotz and Miller (1993) and Hotz, Miller, Sanders, and Smith (1994) bypass the computation of optimal policies (the approach followed in say Rust (1987)'s study of a single agent's dynamic optimization problem) by estimating policies directly from the choices that agents make. This idea has been adapted to strategic settings by several recent papers in Industrial Organization such as Bajari, Benkard, and Levin (2007), Pakes, Berry, and Ostrovsky (2007), Pesendorfer and Schmidt-Dengler (2008), Ryan (2006) and Dunne, Klimek, Roberts, and Xu (2006). Specifically, if $\hat{\Psi}[a_i|x]$ and $\hat{D}$ is available from the data, then the W function can be computed using the DASA up to a vector of parameters in the period profit function. The main assumption required for these techniques to work is that there is a single equilibrium played in each observed state (to the econometrician), which allows me to consistently estimate the conditional choice probabilities, i.e. the probability that a firm chooses an action a_i in an observed state x. I employ this approach along with techniques that allow for persistent unobserved heterogeneity between markets.

First I will present the state space, since the bottleneck in any of these approaches remains the complexity of computing the counterfactual equilibria; more precisely, the burden of keeping the entire state space in memory. Second, I will talk about the parametrized profit function that I estimate. Third, I discuss the conditions on the unobserved states under which I can estimate the conditional choice probabilities. Finally, I will present the estimation criterion, the Indirect Inference Conditional Choice Probability estimator (henceforth IICCP).

5.1 State Space

I choose a maximum of 10 plants per market, since this allows me to pick up most counties in the U.S. (note that 6 plants is the 95th percentile of the number of plants in a county in Table 4), and keeps the size of the state space manageable. A county with more than 10 active plants at some point its history is dropped from the sample, since the model does not allow firms to envisage an environment with more than 9 competitors.[19]

Since the vast majority of plants are owned by single-plant firms, I assume that a firm can operate at most one ready-mix concrete plant. Firm i can be described by a firm specific state $s_i^t \in S_i$:

$$s_i^t = \{ \underbrace{x_i^t}_{\text{(Plant Size,Past Plant Size)}}, \underbrace{m}_{\text{Market Effect}}, \underbrace{\varepsilon_i^t}_{\text{i.i.d. logit shock}} \} \tag{10}$$

I assume that the the unobserved state can be decomposed into ε_{ia}^t which is an i.i.d. logit shock plus a market m level component of the value of taking an action a (which will be discussed in more detail in Section 5.3), and that a firm's observed state x_i^t is based on the number of employees at the firm. A firm is small if it has fewer than 8 employees, medium if it has between 8 and 17 employees and big if there are more than 17 employees. Employment is a better measure of size that capital stock, since it is available for all plants in the data and the number of employees has significantly higher autocorrelation than capital assets and is a better predictor of both future production and exit. [20] I keep track of the largest size the firm has had at any point in the past to control for the fact that a plant which was previously large may have capital assets which make it easier to ramp up its size in the future. Table

[19]To allay the potential for selection bias that this procedure entails, counties with more than 10 000 construction employees at any point between 1976 and 1999 are also dropped. This excludes 15% of markets and 35% of plants from the analysis.

[20]I choose cutoffs of 8 and 18 employees, since these correspond to terciles of the empirical distribution.

Curent Size	Size Next Year				
	Out	Small*	Medium**	Large***	Total
Out	98.7%	1.0%	0.2%	0.1%	14,440
Small	7.9%	82.6%	8.1%	1.5%	1,134
Small, Medium in Past	7.8%	73.7%	17.6%	1.0%	405
Small, Large in Past	10.8%	67.4%	15.8%	6.0%	139
Medium	3.3%	19.9%	68.7%	8.1%	672
Medium, Large in Past	3.0%	10.9%	64.4%	21.6%	298
Large	2.6%	4.1%	11.3%	82.1%	887

* Small Less than 8 Employees
** Medium: 8 to 17 Employees
*** Large: More than 17 Employees

Table 5: Average Yearly Plant Transition Probabilities

5 shows the probability with which plants change size, enter or exit. Large plants exit at a rate of 2.6%, one third the rate of small plants (7.9%), and plants that were large in the past are more likely to ramp up in the future. Moreover, in the sample of counties that excludes large markets 54% of plants are small, 32% are medium and 12% are big. I do not keep track of size in the past if the firm is larger today than it was in the past, since it is the largest size of previous employment that determines if a firm has the equipment and land necessary to ramp up in the future.[21]

The main empirically relevant implication of the assumptions on ε is that it is serially uncorrelated, uncorrelated between players, and privately observed.

[21] Moreover, if I keep track of past size, no matter what current size is, there would be 10 possible common knowledge states x_i^t per firm instead of the 7 in Table 5, which would be a very large increase in the number of states I need to keep track of.

Assumption 1 *(Private Information) Each firm privately observes ε_i^t before choosing its action, a_i^t.*

The assumption that unobservables for the econometrician are also unobserved by other firms in the market is a strong one. Firms typically have detailed information on the operations of their competitors. In contrast to a static model of entry with private information such as Seim (2006), the probability that a "mistake" occurs such as two firms entering into a market at the same time is quite small since entry and exit rates are only 6%, and firms have the option to exit in the next period.

Assumption 2 *(No Market Level Correlation) Unobserved states are independent across firms within the same market, i.e. $\varepsilon_i^t \perp \varepsilon_{-i}^t$.*

This assumption rules out common shocks, or persistent market level unobservables. This assumption will be problematic since there is a clear problem of market unobservables such as either higher costs in certain markets rather than others, such as unionized workers in Illinois but not in Alabama, or higher unobserved demand in some markets, such as the fact that asphalt- but not concrete- melts on roads in Texas. To correct for this problem I include market effects m into firm's profit function, and I will estimate a profit function $r^m(a_i^t, x^t|\theta)$ which differs by market. Note that this implies that the firm's choice specific value functions $W^m(a_i, x)$ and policy functions $\Psi^m(a_i|x)$ will also be market specific. While estimating the market level profit function is straightforward, it runs into serious data constraints, since I cannot identify parameters from the cross-section. Thus I will make two assumptions to render the market effects m tractable. First, I assume that rewards in a market m are additively separable in the market level component:

$$r^m(a_i^t, x^t|\theta) = r(a_i^t, x^t|\theta) + \xi_a^m + \varepsilon_{a_i^t} \tag{11}$$

and have a market/action effect $\xi_{a_i}^m$. Second, I collapse market effects into market category effects μ. Categories are constructed by rounding the average number of plants in a county to the nearest integer. This classification scheme conditions on an endogenous variable, but in appendix A I show that using similar classification schemes that are not endogenous (but harder to fit into the model), such as classifying based on the number of plants in the past, yields indistinguishable estimates. Table 6 shows multinomial logits of a firm's choice of it's size next year (a_i^t), where Column II has market category effects, and Column I does not. Later in the paper, I refer to the predictions using the estimates in Column I as $\hat{P}$ and using Column II as $\hat{P}^\mu$. Note that introducing market effects leads to significantly more negative effects of competitors, where the effect of more than 1 competitor is positive without market category effects, but turn negative with category effects. As well, the effect of log country construction employment falls when market effects are added. In appendix A I discuss the use of the average number of firms in the 1976-1999 periods as proxy for including market fixed effects by presenting binary logit regressions of the decision to have an active plant. Market category controls such as the number of firms in a pre-period or average log construction employment yield similar results to having no market controls, while using either the average number of plants or the average number of plants in all years prior to this year yield similar results to market fixed effects, estimated via a conditional logit.

If the market level shock is ignored then the implicit unobservable is $\tilde{\varepsilon}_i^t = \xi_a^m + \varepsilon_{a_i^t}$. Firms will react to ξ_a^m, since higher values of the market level shock will be more profitable markets to enter, and thus the number of firms is correlated with the observable, i.e. $E[\tilde{\varepsilon}_i^t N^t]$. This will lead to upward bias in the competition coefficient. It turns out that estimating positive effects of competition has a toxic effect on both estimation and counterfactuals, since simulating the model forward with positive spillovers between firms makes the market tip from no firms to being completely filled up with firms. However market level effects wash out a large part of the correlation between demand and the number of firms since much of this correlation is coming from cross-sectional variation. This is a similar problem as the use of fixed effects in a production function regression discussed in Griliches and Mairesse (1998), where fixed effects eliminate the most important source of variation in capital stock, leading to a downward bias on the capital coefficient. These two biases are a serious problem since having no market level controls will lead to a market where the number of firms sloshes around since competition effects are too small to pin down the number of firms, while a model with market fixed effects will predict too small a response to demand and too few changes in the number of plants.

Assumption 3 *(Serial Independence) Unobserved states are serially independent, i.e.* $\Pr(\varepsilon_i^t|\varepsilon_i^k) = \Pr(\varepsilon_i^t)$ *for* $k \neq t$.

Serial independence of unobserved components of a firm's profitability is violated by any form of persistent productivity difference between firms, or long term reputations of ready-mix concrete operators. Note that in the context of a dynamic game, unobserved states are a first-order problem since the size of the firm-level state x_i^t is severely restricted by the difficulty of keeping track of the joint distribution of the states of all firms.

I simulate the age profile of exit using the exit and size changes in Table 5, which captures what the age profile of exit would look like in the absence of selection on an unobserved state. With a serially correlated unobserved state, as plants age their exit rate falls due to the effect of selecting out plants with a bad unobserved state. Figure 3 shows the exit hazard with age in the data and simulated data. Both the data and the simulation have the same average exit rate of about 6%, but the data has a somewhat steeper decline in exit rates over time, so a plant aged 20 years old has an exit rate of about 3.5% in the data, while the simulated data yields a exit rate of about 5.2%. This is consistent with most models of industry dynamics with a serially correlated unobserved state, and the active or passive learning models of Pakes and Ericson (1998) and Jovanovic (1982), but is a small effect compared to other industries such as restaurants where we would worry more

Dependent Variable	Independent Variable	I Coeff.	I S.E.	II (Market Category) Coef.	II (Market Category) S.E.
Small	Small	6.75	(0.03)	6.57	(0.03)
in $t+1$	Small, Medium in Past	6.58	(0.04)	6.31	(0.04)
	Small, Large in Past	6.13	(0.06)	5.90	(0.06)
	Medium	6.15	(0.05)	5.93	(0.05)
	Medium, Large in Past	5.59	(0.08)	5.31	(0.08)
	Large	4.74	(0.06)	4.51	(0.06)
	Log County Employment	0.12	(0.01)	-0.05	(0.01)
	First Competitor	-1.65	(0.05)	-1.96	(0.05)
	Second Competitor	0.06	(0.03)	-0.47	(0.03)
	Third Competitor	0.11	(0.04)	-0.29	(0.04)
	Log of Competitors above 3	0.13	(0.02)	-0.03	(0.03)
	Market Type 2			1.03	(0.03)
	Market Type 3			1.65	(0.05)
	Market Type 4			2.26	(0.06)
	Constant	-3.82	(0.07)	-3.02	(0.07)
Medium	Small	6.27	(0.05)	6.10	(0.05)
in $t+1$	Small, Medium in Past	6.95	(0.06)	6.68	(0.06)
	Small, Large in Past	6.43	(0.08)	6.21	(0.08)
	Medium	9.17	(0.06)	8.96	(0.06)
	Medium, Large in Past	9.12	(0.08)	8.84	(0.08)
	Large	7.47	(0.06)	7.25	(0.06)
	Log County Employment	0.29	(0.01)	0.12	(0.01)
	First Competitor	-1.76	(0.05)	-2.11	(0.05)
	Second Competitor	-0.03	(0.04)	-0.54	(0.04)
	Third Competitor	0.04	(0.05)	-0.32	(0.05)
	Log of Competitors above 3	0.02	(0.03)	-0.11	(0.03)
	Market Type 2			1.08	(0.04)
	Market Type 3			1.68	(0.06)
	Market Type 4			2.21	(0.07)
	Constant	-6.47	(0.09)	-5.71	(0.09)
Large	Small	5.04	(0.08)	4.88	(0.08)
in $t+1$	Small, Medium in Past	4.50	(0.12)	4.23	(0.12)
	Small, Large in Past	5.81	(0.10)	5.59	(0.10)
	Medium	7.46	(0.07)	7.26	(0.07)
	Medium, Large in Past	8.39	(0.09)	8.13	(0.09)
	Large	9.76	(0.07)	9.54	(0.07)
	Log County Employment	0.51	(0.01)	0.34	(0.02)
	First Competitor	-1.81	(0.06)	-2.17	(0.06)
	Second Competitor	-0.05	(0.05)	-0.58	(0.05)
	Third Competitor	-0.02	(0.06)	-0.39	(0.06)
	Log of Competitors above 3	-0.02	(0.03)	-0.15	(0.03)
	Market Type 2			1.07	(0.06)
	Market Type 3			1.71	(0.07)
	Market Type 4			2.26	(0.08)
	Constant	-8.28	(0.11)	-7.50	(0.11)
	Observations	431420		431420	
	Log-Likelihood	-87256		-86302	
	Likelihood Ratio	431431		433339	

Table 6: Multinomial Logit on the choice to be large, medium or small.

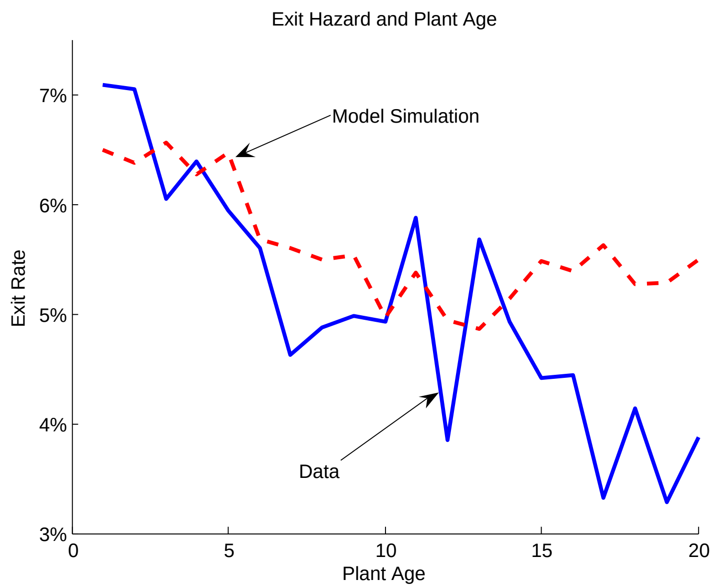


Figure 3: The data predicts a slightly steeper decline of the exit hazard with age.

about unobserved states. I do not deal with serial correlation and both the estimates and counterfactuals will be contaminated by this problem.

5.2 Profit Function

The reward function has parameters, θ, which will be recovered from the data. I use a simple Bresnahan and Reiss (1991) style reduced-form for the reward function, endowed with parameters θ. It is easily interpreted and separable in dynamic parameters:

$$r(a_i^t, x^t|\theta) = \sum_{\alpha \in \{\text{Big,Medium,Small}\}} 1(a_i^t = \alpha) \times \left(\underbrace{\theta_1^\alpha}_{\text{Fixed Cost}} + \underbrace{\theta_2^\alpha M^{t+1}}_{\text{Demand Shifter}} \underbrace{+ \theta_3^\alpha g(\sum_{-i} x_{-i}^{t+1} \neq \text{out})}_{\text{Competition Parameters}} \right) \tag{12}$$

where $g(\cdot)$ is a non-parametric function of the number of competitors.

Transition costs are:

$$\tau(a_i^t, x_i^t|\theta) = \theta_4^{l,m} \sum_{l>0, m \neq l} 1(a_i^t = l, x_i^t = m) \tag{13}$$

so a firm pays a transition cost to change its state. However, I assume that a firm does not pay any exit costs. [22]

5.3 Indirect Inference CCP Algorithm

Suppose I estimate the model by matching the optimal choice probabilities $\Psi(a_i|x, W^\theta, \theta)$ to the data, where I include θ to emphasize the fact that the choice probability Ψ and choice specific value function W depend on a parameter vector θ. The natural way to do this would be to compute an equilibrium to the dynamic game given parameters θ and the choice specific value function W^θ associated with it. Computing a solution to the dynamic

[22] Monte Carlo experiments indicate that it is quite difficult to jointly identify fixed costs, entry and exit costs.

game for each candidate parameter vector θ is computationally impractical, since even with the DASA it takes more than an hour to find an equilibrium. Moreover, the choice probabilities $\Psi(a_i|x, W^\theta, \theta)$ might be quite non-linear in θ since the solution of the dynamic game W^θ can have a very intricate shape.

To cut through this difficult dynamic programming problem, I have adapted a conditional choice probability estimator applied to games. I have adapted the CCP algorithm for the very large state space in this problem (over 350000 states) using the Stochastic Algorithm of Pakes and McGuire (2001) and I use a Simulated Indirect Inference Criterion approach for estimation (Keane and Smith Jr (2003),Gourieroux and Montfort (1993) and Gourieroux and Monfort (1996)).[23]

Algorithm CCP Indirect Inference Algorithm (CCPII)

1. Replace optimal choice probabilities Ψ with an estimate from the data $\hat{P}^\mu$. Estimate the demand transition process $\hat{D}^\mu[M'|M]$.

   I assume that there is single symmetric Markov Perfect equilibrium played in each market category μ, and thus each market category μ will have CCPs $P^\mu = \{\Pr(a_i^t|x^t)\}_{a_i^t, x^t}$ associated with it.[24] Since the

[23] In previous version I have computed present estimates using an approach in the spirit of Aguiregabiria and Mira (2007) which iteratively updates the strategies used by firms, and I find that using an iterated technique yields very similar results to those presented in the paper.

[24] To compute counterfactual industry dynamics I assume the existence of a single symmetric Markov perfect equilibria per market category μ. Besanko, Doraszelski, Kryukov, and Satterthwaite (2010) show that this assumption may be problematic.

ϵ_i^t's are logit draws, I estimate $\hat{P}^\mu$ with a multinomial logit that varies by market category μ shown in Table 6. Due to limited data, rather than estimating coefficients on the logit $\beta_{a_i}^{\mu,0} + \beta_{a_i}^{\mu,X} X$ that all vary by market category, I assume that the market effects are just additive constants, i.e. $\beta_{a_i}^{\mu,0} + \beta_{a_i}^{X} X$.[25]

The demand transition matrix D is estimated by market category μ using a bin estimator $\hat{D}^\mu[i|j] = \frac{\sum_{(l,t)} 1(M_l^{t+1} \in B_i, M_l^t \in B_j)}{\sum_{(l,t)} 1(M_l^t \in B_j)}$ with 10 bins and the demand level within a bin is set to the mean demand level.

2. Compute the W^μ function conditional on policies $\Psi^\mu(a_i|x, W) = \hat{P}^\mu[a_i|x]$ using the DASA.

   The choice specific value function W^μ can be computed using the DASA with an important change. I will replace the computed choice probabilities $\Psi^\mu(a_i|x, W^\mu)$ with an estimate from the data $\hat{P}^\mu[a_i|x, W^\mu]$. The considerably simplifies the DASA, since it means that I don't have to compute the choice probabilities Ψ, since these are being held fixed during the iteration of the algorithm. [26]

   A final rewriting of the W-function is now in order to aid with the estimation of the model. The rewards and transition costs in equations (12) and (13) on page 36 are linear in parameters θ, so the profit func-

[25] The main issue is that it is difficult to estimate the effect of say the 3rd competitor, in a market that has on average one firm in it-hence in market category $\mu = 1$, since we rarely see 3 firms in this type of market.

[26] I modify the DASA replacing $\Psi^\mu(a_i|x, W^\mu)$ with $\hat{P}^\mu[a_i|x, W^\mu]$, and I shut down the policy update step; i.e. step 6 on page 25.

tion can be rewritten as $r(a_i, x|\theta) - \tau(a_i, x_i|\theta) = \theta \cdot \vec{\rho}(a_i, x)$ where $\vec{\rho}$ is a function that returns a vector. This implies that the W function will be separable in dynamic parameters as in Bajari, Benkard, and Levin (2007), since

$$\begin{aligned} W^\mu(a_i, x|\theta) &= \mathbb{E}\sum_{t=1}^{\infty}\beta^t\left(r(a_{it}, x_t|\theta) - \tau(a_{it}, x_{it}|\theta)\right) \\ &= \theta \cdot \mathbb{E}^{P^\mu}\sum_{t=1}^{\infty}\beta^t\vec{\rho}(a_{it}, x_t) \equiv \theta \cdot \Gamma^{P^\mu}(a_i, x) \end{aligned} \tag{14}$$

where $\mathbb{E}^{P^\mu}$ denotes the expectations of firms at time 0 given that they believe their opponents and their future selves use conditional choice probabilities P^μ. This allows me to rewrite the W^μ function in a vector representation, the Γ^{P^μ} function which only depends on the conditional choice probabilities P^μ, not on the parameter vector θ:

$$\Gamma^P(a_i, x) = \mathbb{E}^{P^\mu}\sum_{t=1}^{\infty}\beta^t\vec{\rho}(a_{it}, x_t) \tag{15}$$

The choice probabilities Ψ can be rewritten as a function of Γ^{P^μ} and θ:

$$\Psi(a_i|x, \Gamma^{P^\mu}, \theta) = \frac{\exp\left(\theta \cdot \Gamma^{P^\mu}(a_i, x)\right)}{\sum_{j\in A}\exp\left(\theta \cdot \Gamma^{P^\mu}(j, x)\right)} \tag{16}$$

I compute the Γ function with the modified Discrete Action Stochastic Algorithm and the information on the evolution of the state in the estimated choice probabilities $\hat{P}^\mu$ and the estimated demand process

$\hat{D}^{\mu}$. [27]

3. Simulated Indirect Inference Estimation

   A maximum likelihood estimation strategy using the choice probabilities Ψ is quite practical since the log-likelihood function $\mathcal{L}$ will be globally concave[28]:

$$\mathcal{L}(\theta) = \sum_{n=1}^{N} \ln\left(\Psi(a_i^n | x^n, \theta, \Gamma)\right) \tag{18}$$

   where n indexes observations in the data. A problem arises because Γ has error in it, both because the choice probabilities $\hat{P}^{\mu}$ have sampling error and because the discrete action stochastic algorithm is an approximation to the true value function. With simulation error, maximum likelihood estimates will be biased (see McFadden (1989) and Pakes,

[27] I could have computed the $\Gamma^{P^{\mu}}$ using forward simulation, i.e.:

$$\Gamma^{P^{\mu}}(a_i, s) \approx \frac{1}{K} \sum_{k=1}^{K} \sum_{t=0}^{\infty} \beta^t \vec{\rho}(a_{itk}, x_{tk}) \tag{17}$$

where the sequence of states x^{tk} can be simulated using demand transition process $\hat{D}$ and the choice probabilities for firms $\hat{P}$. However, there are about 350 000 states and 4 actions, thus I would need to do this forward simulation 1.4 million times the number of simulation draws K. To get around this computational burden, I use a modification of the DASA algorithm, in which I shut down the policy update step, i.e. step 6 on page 25, to perform the forward simulation in equation 17 more quickly. In appendix C I discuss the computation of Γ^{P} in more detail.

[28] To see this, note that the assumption of linearity in dynamic parameters gives a utility function of the form $u_a = \theta \cdot \Gamma(a, x)$ which is linear. Along with the assumption that the ϵ_a's are logit, this implies a globally concave likelihood function. I use maximum likelihood estimates as starting values for the indirect inference procedure which are presented in Table 14 on page 79 in the Appendix. The main difference between the estimates using maximum likelihood and those using Indirect Inference, is that maximum likelihood estimates a higher variance of ε than Indirect Inference.

Berry, and Ostrovsky (2007) for instance), and this bias may disappear quite slowly as one increases the precision of Γ. Moreover, maximum likelihood estimates are frequently pinned down by small probability events (such as entry and exit), which are very sensitive to error in the Γ function. Instead, I use an Indirect Inference Criterion function to estimate the model. Indirect Inference is less sensitive to error in the Γ function, and like many GMM estimators, and can be consistent even if there is simulation error in Γ, and this simulation error does not vanish asymptotically.[29]

Essentially the indirect inference estimator matches regression coefficients from the data (denoted $\hat{\beta}$) with regression coefficients from simulated data generated by the model conditional on a parameter θ (denoted $\tilde{\beta}(\theta)$). As an auxiliary model (i.e the regression I run on both real and simulated data) I choose a multinomial linear probability model. It is simple to estimate and a close analogue to the multinomial dynamic logit model which is being estimated.[30]

I define the outcome vector from the data as $\mathbf{y_n}$, and the predicted

[29]For some intuition, if the exit rate in the data is 1%, but the model predicts an exit probability almost 0%, then an maximum likelihood criterion would have an infinite log-likelihood, while a Indirect Inference criterion would find an error of 1%.

[30]Note that the auxiliary model does not need to be a consistent estimator or have an interpretation of any sort. It's sole responsibility is to provide rich description of the patterns of a dataset, and to be simple to estimate.

choice probabilities given by the model $\tilde{\mathbf{y}}_{\mathbf{n}}$ for observation n as:

$$\mathbf{y}_{\mathbf{n}} = \begin{bmatrix} 1(a_n = \text{small}) \\ 1(a_n = \text{medium}) \\ 1(a_n = \text{big}) \end{bmatrix} \qquad \tilde{\mathbf{y}}_{\mathbf{n}}(\theta) = \begin{bmatrix} \Psi(\text{small}|x_n, \Gamma, \theta) \\ \Psi(\text{medium}|x_n, \Gamma, \theta) \\ \Psi(\text{large}|x_n, \Gamma, \theta) \end{bmatrix} \tag{19}$$

Where the outcome vector $\tilde{\mathbf{y}}_{\mathbf{n}}(\theta)$ are the predicted choice probabilities Ψ. [31] I run an OLS regression on $y_n = \mathbf{Z}_n\hat{\beta}$ and find the ols coefficients of the multinomial linear probability model. Likewise I run a OLS regression on the predicted choice probabilities $\tilde{y}_n$ to obtain the coefficients for the model $\tilde{\beta}(\theta)$ given parameter θ. Table 7 shows this multinomial linear probability model, the $\hat{\beta}$ coefficient . The covariates of the auxiliary model z_n are indicators for the firm's current state, the number of competitors in a market and the log of construction employees in the county. I allow the coefficients to vary by market category μ and by action chosen a_n. This tells to the model to match moments conditioned on market category μ to take care of heterogeneity between markets.

The criterion function minimizes the distance between the regression coefficient in the data and in the simulated data:

$$\mathcal{Q}(\theta) = \left(\hat{\beta} - \tilde{\beta}(\theta)\right)' \mathbf{W} \left(\hat{\beta} - \tilde{\beta}(\theta)\right) \tag{20}$$

[31] Theorem 4 in the appendix proves that using the choice probabilities Ψ as predicted actions gives the same θ's as drawing action $a_n \sim \Psi(\cdot|x_n, \Gamma, \theta)$ from the predicted choice probabilities when one uses an infinite number of simulation draws.

	Dependant Variable					
	Small		Medium		Large	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Small	0.801	(0.001)	0.075	(0.001)	0.013	(0.001)
Small, Medium in Past	0.711	(0.002)	0.168	(0.002)	0.007	(0.001)
Small, Large in Past	0.650	(0.003)	0.151	(0.003)	0.056	(0.002)
Medium	0.174	(0.002)	0.679	(0.001)	0.078	(0.001)
Medium, Large in Past	0.084	(0.002)	0.637	(0.002)	0.211	(0.002)
Large	0.018	(0.001)	0.106	(0.001)	0.814	(0.001)
Log County Employment	-0.004	(0.000)	0.001	(0.000)	0.003	(0.000)
First Competitor	-0.043	(0.001)	-0.018	(0.001)	0.003	(0.001)
Second Competitor	-0.005	(0.001)	-0.003	(0.001)	-0.001	(0.001)
Third Competitor	-0.005	(0.001)	-0.001	(0.001)	-0.003	(0.001)
Log of Competitors above 3	0.001	(0.001)	-0.001	(0.001)	-0.002	(0.001)
Market Type 2	0.013	(0.001)	0.004	(0.001)	-0.001	(0.001)
Market Type 3	0.024	(0.001)	0.005	(0.001)	0.001	(0.001)
Market Type 4	0.037	(0.002)	0.005	(0.001)	0.008	(0.001)
Constant	0.068	(0.002)	0.014	(0.002)	-0.020	(0.001)
Observations	431420		431420		431420	
R^2	62%		47%		68%	
F-Stat	49496		26888		64486	

Table 7: Multinomial Linear Probability Model of Plant Size

where $\mathbf{W}$ is a weighting matrix, where I use $\mathbf{W} = Var[\hat{\beta}]^{-1}$, the inverse of the covariance matrix from the OLS regression. Appendix D.1 shows conditions under which the estimator is consistent, which a straightforward extension of the consistency of Indirect Inference estimators.[32]

To recap, IICCP estimation proceeds in three stages. First, I estimate the

[32]In a prior version of the paper I estimated the model by iterated on the conditional choice probabilities, i.e. updating them using parameters estimates θ. To implement this procedure (which requires the assumption of a single equilibrium to the dynamic game in order to be consistent) I need to add extra steps where:

4) Replace $\hat{P}^{\mu}[a_i|x]$ with $\Psi[a_i|, x, \hat{\theta}]$ where $\hat{\theta}$ is the current estimate of the parameters in the profit function.

5) Repeat steps 2-4 until θ converges.

I obtain very similar results when I iterate on the conditional choice policies as when I do not.

policies used by firms $\hat{P}^{\mu}$ and the demand transition process $\hat{D}^{\mu}$. Second, I use these estimates - that let me simulate the evolution of the state - and the DASA to compute the choice specific value function W up to a vector of parameters in the period profit function θ. This allows me to quickly compute a firm's policy function $\Psi(a_i|x, \Gamma, \theta)$. Third, I pick θ so that the coefficients in an OLS regression using data simulated by the model are as close as possible to the OLS coefficients using the actual data.

6 Results

I estimate the model using the CCP Indirect Inference presented in the previous section. I fix the discount factor to 5% per year. Table 8 presents estimates of the dynamic model using the CCP Indirect Inference. Column I shows estimates without fixed costs that vary by market category $\hat{\theta}$, while Column II shows estimates where fixed costs vary by market category $\hat{\theta}^{\mu}$. Standard errors are computed via 100 bootstrap replications of the estimation procedure, where I reestimate the demand transition process $\hat{D}^{\mu}$ and the conditional choice probabilities $\hat{P}^{\mu}$, then minimize the criterion function $\mathcal{Q}$ to find θ. I block bootstrap by market, resampling a market's history from 1976 to 1999, so the computed standard errors account for serial correlation within a market.

In line with interviews with producers in Illinois, I calibrate the entry costs for a medium sized plant to 2 million dollars. This allows me to convert parameters in variance units into dollars. The variance of ε is estimated to 120 000 dollars per year (or \$ 97 000 for the market effects model), or about 4% of sales which is well below year to year changes in profits due to changes in productivity. To make sense of the magnitudes of these figures, note that average sales are 3.4 million dollars.

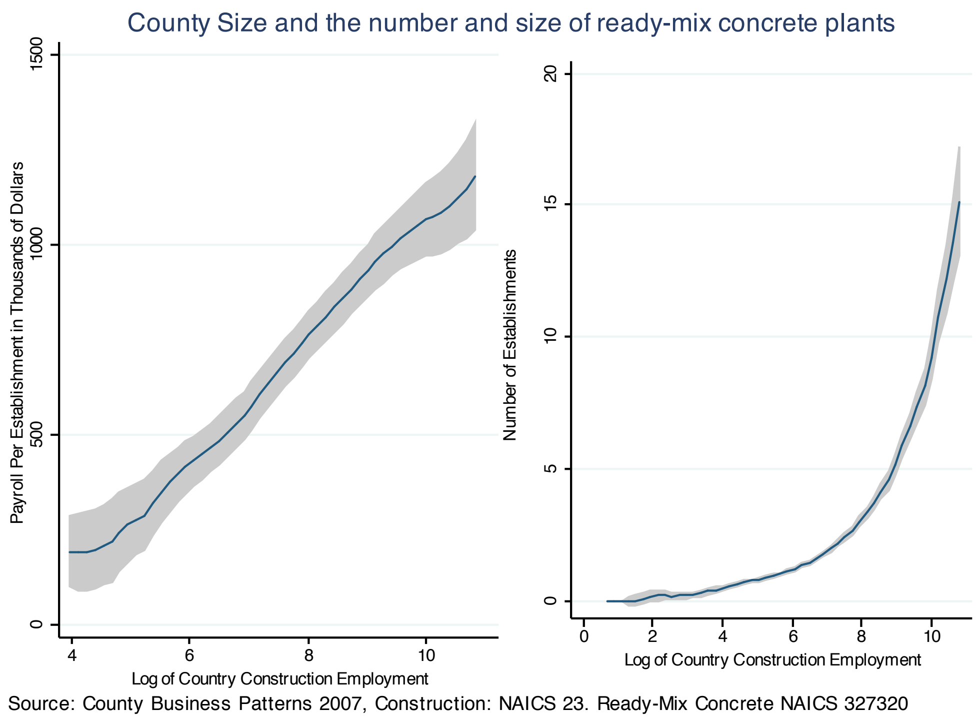
The fixed costs of operating a plant are about \$ 261 000 for a medium sized plant, slightly less for a small plant and slightly more for a large plant. Doubling the number of construction workers in a county increases profits by \$ 12 000 for a small plant versus \$ 44 000 and \$ 83 000 for a medium and large sized plant, while the effect of construction employment is somewhat lower in the market effect estimates $\hat{\theta}^{\mu}$. The coefficient on construction employment is reflective of the fact that bigger markets have both more plants and larger plants. Figure 4 plots local polynomial regressions of the log of construction employment in the county against both the average size of a plant (measured in payroll terms) and the number of plants in a county: larger markets have more plants and bigger plants. In a county with 150 employees in the construction sector average plant payroll plants is \$ 400 000, while in a county with 1010 employees in the construction sector this average is closer to \$ 600 000. This effect is not specific to the ready-mix concrete industry, as Campbell and Hopenhayn (2005) have documented the link between establishment and market size in retail trade industries. The fact that plant size and market size are connected is responsible for the fact that demand fluctuations have a impact on the equilibrium size distribution. In

(All estimates in thousands of dollars)

		I. No Market Effects $\hat{\theta}$		II. Market Effects $\hat{\theta}^{\mu}$	
		Coef.	S.E.*	Coef.	S.E.
Fixed Cost	Small	-237	(11)	-246	(13)
	Medium	-261	(11)	-291	(16)
	Large	-287	(12)	-331	(18)
Log Construction	Small	12	(8)	-4	(6)
Employment	Medium	44	(6)	29	(6)
	Large	83	(7)	65	(7)
1st Competitor	Small	-96	(9)	-170	(12)
	Medium	-101	(12)	-179	(13)
	Large	-76	(13)	-148	(16)
Log Competitors	Small	-73	(7)	-48	(7)
(above 1)	Medium	-67	(8)	-45	(7)
	Large	-40	(9)	-32	(11)
Market Effects					
Category 2				200	(9)
Category 3				55	(6)
Category 4				92	(7)
Transition Costs					
Out →Small		-918	(37)	-1301	(24)
Out →Medium †		*-2000*	(26)	*-2000*	(44)
Out →Large		-2721	(28)	-2907	(37)
Small →Medium		-528	(26)	-588	(27)
Small, Past Medium →Medium		-824	(31)	-879	(33)
Small, Past Large →Medium		-299	(25)	-270	(19)
Small →Large		-1973	(26)	-1973	(30)
Small, Past Medium →Large		-460	(29)	-386	(27)
Small, Past Large →Large		-277	(26)	-174	(11)
Medium →Small		-7	(12)	32	(13)
Medium, Past Large →Small		-289	(20)	-198	(25)
Medium →Large		105	(22)	60	(22)
Medium, Past Large →Large		30	(23)	15	(22)
Large →Small		-111	(18)	-179	(22)
Large →Medium		-427	(21)	-459	(24)
Standard Deviation of Shock		120		97	
GMM Criterion		10958		10822	

†: The entry costs of a medium sized plant are calibrated to 2 million dollars.
∗: Standard Errors are computed using 100 block bootstraps.

Table 8: Estimates for the Dynamic Model of Entry, Exit and Investment.



Note: Local Polynomial Regressions with shaded areas representing the 95% confidence interval.

Figure 4: Larger Markets have both more plants and larger plants.

particular due to the linearity of the relationship between the size of establishments and the log of construction employment, lower variance of county construction employment (with the same mean) will raise the size of establishments. Furthermore, regressing the log of the number of establishments on log construction employment, I find a coefficient of 0.69, so a 1% increase in construction employment increases the number of firms by less than 1%. This implies that the number of plants in county is a concave function of construction employment for the range of small markets I am looking at.

The number of competitors in the county has large effect on profits. The first competitor reduces profits by $ 101 000 for a medium sized plant, and doubling the number of competitors (beyond the first competitor) reduces profits by $ 67 000 per year. Note that the effect of the first competitor is larger than the effect of subsequent plants, which echoes the Bertrand like nature of competition in the industry. When I include market fixed effects, I find somewhat more negative effects of competition for the first competitor and somewhat smaller effects for subsequent competitors. [33]

The patterns in the transition costs are also illuminating and reflect the transition patterns for plant size found in Table 5 on page 29. Entry costs are $ 0.9 million for small plants and $ 2.7 million for large plants. This is in line with substantial differences in machinery and land for larger plants. There are also large costs of increasing the size of a plant, of about $ 0.5 million to grow a plant from small to medium, $ 2.0 million to get it from small to large, and $ 0.1 million to ramp a plant from medium to large. Note that it is cheaper to enter as a small plant and grow to a large plant in the next period and 80% of plants enter as small plants. Finally, the model also estimates substantial costs of ramping the size of plant back down. These large transition costs imply that plants will have a fairly weak response to demand shocks on either the extensive (or entry) margin or on the intensive

[33] Note that if I remove market indicators from the covariates z in the auxiliary regression shown in Table 7, I find substantially smaller effects of competition, so it is still important to target these market level moments, even if they are not captured by changes in market fixed costs.

(or size) margin. For instance for a medium sized plant to recover a sunk entry cost of 2 million dollars would require an increase in demand of two log points in perpetuity, since the coefficient on demand is \$ 44000 and projecting it forever yields $\frac{44{,}000}{\beta} = 880,000$.

A bigger size in the past reduces the costs of growing a plant. Small plants that were medium in the past or large in the past find it easier to ramp up their size (with the exception of small plants who were medium becoming large plants). Likewise a medium sized plant that was large in the past has a lower cost of reverting back to being a large plant. The dependence of transition costs on size in previous years lowers the implicit adjustment costs, since a plant can shrink today and retain the ability to cheaply increase it's size in the future.

6.1 Model Fit

To evaluate the fit of the model, I compare the evolution of the concrete market from 1976 to 1999 to the evolution predicted by the model. To simulate the model's predictions, I use the discrete action stochastic algorithm (DASA) to compute an equilibrium to the dynamic game using both the no market fixed effect (henceforth $\hat{\theta}$) and market fixed effect (henceforth $\hat{\theta}^{\mu}$) parameters estimates in Table 8. Note that this equilibrium needs to be computed for all 4 market categories since they have a different demand $\hat{D}^{\mu}$ process, and will also have different fixed cost when I use the market fixed effect parameter estimates $\hat{\theta}^{\mu}$. Using the computed policies and demand transition process, the model is simulated until 1999, where markets in 1976 are used as initial states x^0. The CCPII algorithm used year to year moments to estimate parameters rather than predictions on the entire path of the industry over time, so the path of the industry generated by the model's prediction could in principle differ substantially from the path of the industry found in the data. Table 9 shows moments of the actual distribution of firms, and the simulated evolution of firms for both parameters without ($\hat{\theta}$) and with ($\hat{\theta}^{\mu}$) market varying fixed costs.

I will focus on two types of moments: plant level moments and market level moments. Both simulations do well at matching the distribution of plant size, with 50% of small plants in the data and in the simulations, and 24% large plants in the data versus 33% and 29% in the simulated data with and without market fixed effects. As for entry and exit rates, these are about 7.5% in the data, and 6.8% in the simulated data without market fixed effects ($\hat{\theta}$), but are 25% in the model with market fixed effects ($\hat{\theta}^{\mu}$). The model without fixed effects $\hat{\theta}$ predicts than 10% of plants will increase their size, and 9% will reduce their size, which matches the rate at which plants grow and shrink in the data. However, the model with market fixed effects $\hat{\theta}^{\mu}$ predicts twice the rate of size changes. Overall, the model without fixed effects does a good job at matching plant level dynamics as well as the size composition of the industry.

The second set of moments looks at how well the model predicts the number of plants in a market and the variation of the number of plants

Moments	Real Data (1976-1999)	Simulated Data $\hat{\theta}$	Simulated Data $\hat{\theta}^{\mu}$ with Market Category Effects
Plant Level Moments			
Share of Small Plants	50%	50%	50%
Share of Medium Plants	26%	17%	21%
Share of Large Plants	24%	33%	29%
Entry/Exit Rate	7.5%	6.8%	25%
Ramping Up Rate	10%	10%	19%
Ramping Down Rate	9%	9%	23%
Market Level Moments			
Number of Plants Per Market	2.0	1.8	2.7
No Plants in Market	1%	2%	45%
Monopoly Market	45%	61%	9%
Duopoly	27%	21%	4%
More than 2 plants	27%	15%	41%
Number of Plants in Category 1	1.14	1.00	1.27
Number of Plants in Category 2	1.87	1.64	2.84
Number of Plants in Category 3	2.78	2.45	4.29
Number of Plants in Category 4	4.40	4.00	5.16
Coefficient of Variation			
Number of Plants within Market	0.5	0.5	1.2
Correlation Demand and Plant Size	0.23	0.22	0.25
Correlation Demand and Number of Plants	0.53	0.43	0.39

Note: Both models were estimated using the market category effect conditional choice probabilities $\hat{P}^{\mu}$ to simulate the evolution of the state over time. They differ due to the inclusion or not of market category profit shifters.

Table 9: Model Fit

over time. On average there are 2 plants in the market, while the no market effects model $\hat{\theta}$ forecasts 1.8 plants per market, and the market effects model $\hat{\theta}^{\mu}$ forecasts 2.7. Note that this is the main issue with the market effects model: it predicts a large increase in the number of plants in the market over time. Moreover this large increase in the number of plants is not uniform, as the model with market fixed effects $\hat{\theta}^{\mu}$ predicts than 45% of markets will have no plants and 41% will have more than 2 plants. In contrast, in the data there are 1% of markets with no plants, 45% monopoly markets, 27% duopoly markets, and 27% of markets with more than 2 plants, and model without market fixed effects $\hat{\theta}$ predicts a distribution that is more similar to the distribution in the data. I also show the number of plants in each market category μ, both in the data and in the simulated data. Note that the model without fixed effects $\hat{\theta}$ does a good job at matching the number of plants in each market category, even though the only way that market categories matter is through differences in the estimated demand transition process $\hat{D}^{\mu}$ and the grid of demand states.

To give an idea of how well the model predicts changes in the number of plants, I compute the coefficient of variation (henceforth CV) of the number of plants within a market. The data and the model without fixed effects $\hat{\theta}$ predicts a CV of 0.5, but the model with market fixed effects $\hat{\theta}^{\mu}$ predicts substantially more variation with a CV of 1.2. The correlation between market size and the number of firms is 0.5 in the data, but 0.4 in both models predictions. Likewise, the correlation between market size and plant size (where plant size is just the integers 1, 2 and 3) is about 0.23 which is well matched by either the model without (0.22) and with fixed effects (0.25).

Since the model with market fixed effects fails to match many of the moments on the evolution of the market, I will use the no market effect model to perform counterfactuals. [34]

7 Counterfactual Industry Dynamics

There are substantial local fluctuations in construction activity. How do these demand shocks affect the the ready-mix concrete industry? The counterfactual that I consider would remove much of the short-term fluctuation in construction activity at the county level. There are two margins that are important in evaluating the effect of the demand smoothing policy. The first margin is the dynamic effect of this policy on entry and exit rates and firms ramping up and down their size. These effects may be important since the estimates in Table 8 showed substantial transition costs and thus transitions are costly to plants. Removing fluctuations may have static effects, since it changes the risk that firms are exposed to and the option value of staying in the market. Thus I will also look at the effect of the demand smoothing policy on the size distribution and market structure of the ready-mix concrete

[34] It can be difficult to match the path of the industry over time since small errors in the transition probabilities will accumulate. To illustrate this point I have simulated the evolution of markets using the CCP estimates $\hat{P}^{\mu}$. I find the the "reduced-form" CCP model does substantially worse at predicting the evolution of the market than the market fixed effect model.

industry.

Consider the policy where local governments allocate construction budgets to smooth out changes in demand. The ideal policy would involve sequencing government contracts over a short, say 5 year, period in such a way to minimize the variance of demand. This policy would be fairly easy to implement, since it simply relies on local governments being able to borrow and save over relatively short periods of time, but relies on the fact that construction projects such as roads can be efficiently broken up across years. This exact policy is very difficult to simulate since it involves firms knowing the entire stream of future demand and data on the exact composition of private and government construction activity in each county in each year. Instead, I will simulate the effect of giving firms the expected level of demand given the current demand level over the next 5 years to approximate the idea of getting rid of short-term movements in demand. After 5 years are up, demand reverts to the level it would have had absent of demand smoothing. Thus the long-run path of demand remains unchanged, all that this policy accomplishes is to eliminate short-run wiggles in demand.

Denote the long-run demand level as $\tilde{M}^t$ which represents the demand level at the start of each 5 year period over which demand is smoothed out. Table 10 shows the two demand processes that I consider, the estimated demand process as well as the demand process where firms receive the expected level of demand for the next 5 years. Note that both of these demand processes yield the same long-run demand process, i.e. the expected demand level more than 5 years in the future is exactly the same.

1. Un-smoothed Demand (**Baseline**)

$$M^t \sim \hat{D}[\cdot|M^{t-1}]$$

2. 5 Years of Smoothed Demand (**Policy Counterfactual**)

$$\begin{aligned}\tilde{M}' \sim & \hat{D}^5[\cdot|\tilde{M}] \text{ in periods } t = 5 \text{, set } t = 0 \\ M^t = & \frac{1}{5}\sum_{\tau=1}^{5} E_{\hat{D}(\cdot)}[M^\tau|\tilde{M}]\end{aligned}$$

3. Constant Demand

$$M^t = M^0$$

4. Firms believe demand is constant

Firms Believe	Actual Process
$M^t = M^0$	$M^t \sim \hat{D}[\cdot\|M^{t-1}]$

Table 10: Counterfactual Demand Processes

Equilibrium responses in dynamic oligopoly models are frequently quite difficult to interpret. Thus I show the effect of two other demand smoothing policies for illustrative purposes. The first is constant demand, i.e. $M^t = M^0$ for all t, which illustrates the maximal effect of demand smoothing policies. However, we also want to separate the effect of demand smoothing due to changes in the equilibrium strategies used by firms, such as how responsive they are to demand shocks, versus the direct effect of demand changes holding strategies fixed. To do this, I will also consider the following "myopic" firms, who believe that demand is constant over time, but in fact demand evolves following the estimated process in the data $\hat{D}^{\mu}$.

7.1 Dynamic Effects of Demand Smoothing

Table 11 shows descriptive statistics of the dynamics of the ready-mix concrete industry for the 4 different types demand processes, where I present statistics 25 years after the policy has been put into place to allow the industry to adjust to the new demand process.

	Un-smoothed Demand	5 Years of Smoothing	Constant Demand	Firms believe demand is constant
Turnover				
Exit Rate	3.7%	4.1%	3.6%	6.0%
Change in Size Rate	14%	11%	11%	16%
Investment				
Sunk Entry Costs per year(in million $)	112	206	124	138
Size Changing Costs per year (in million $)	187	308	172	181
Total Plants	2,691	4,339	2,984	2,213

Table 11: Dynamic Effects of Demand Smoothing Policies

Note that there is very little change in the dynamics of the ready-mix concrete industry when the 5 year demand smoothing policy is put into place. The turnover rate stays at 4.1% versus 3.7% in the base case, and the rate at which firms change their size is a somewhat smaller 11% versus 14% in the base case. Moreover, even when *all* demand changes are eliminated, the turnover rate and the size change rate barely change.

The fact that meddling with the demand process has little effect on turnover is consistent with the descriptive work of Dunne, Roberts, and Samuelson (1988), and the fact that the entry and exit rates (per incumbent plant) are virtually uncorrelated at the county-year level.[35] This means that turnover is not generated by market level shocks, which would lead to either entry or exit, but not both, but by idiosyncratic shocks ε_{it}. Yet, this explanation is incomplete since it suggests that demand changes have little effect on firm profits. What happens is that demand fluctuations are anticipated by firms. Thus firms lower their reaction to demand shocks when there is more demand volatility. Notice that if I take firms which use the policies corresponding to a constant level of demand, but subject them to the demand process estimated in the data $\hat{D}^m$, I find that the turnover rate would increase by 50% to 6% per year, and the rate at which firms change their size would go up to 16% per year. Thus expectations of future demand changes are blunting how much firms react to current demand shocks, and this is why we see such a small reaction of turnover and investment to demand changes.

I quantify the effects of industry dynamics on total expenditures on transition costs using the estimates from the dynamic model in Table 8. I find that in the base case, sunk entry costs are $ 112 million dollars per year, while transition costs are $ 187 million per year. When the five year demand smoothing policy is put into place, these costs rise by to $ 206 million per year of sunk entry costs and $ 308 million per year of size changing costs.

[35] Dunne, Roberts, and Samuelson (1988) show that entry and exit rates are highly correlated at the industry level, while I show that in the ready-mix concrete industry are uncorrelated at the county-year level.

This 52% increase in investment is almost entirely due to the 47% increase in the number of plants in the industry due to the 5 year demand smoothing policy. Thus the increase in investment in this industry is due to the static effect of the demand smoothing policy which I now turn to.

7.2 Static Effects of Demand Smoothing

Table 12 shows the static effects of the demand smoothing policy: the distribution of firm size in the industry, the total number of firms, fixed costs expenditures and the market structure of the industry.

	Un-smoothed Demand	Constant Demand	5 Years of Smoothing
Industry Composition			
Small Plants	51%	35%	47%
Medium Plants	16%	8%	5%
Big Plants	31%	57%	47%
Total Plants	2,691	2,984	4,339
Fixed Costs	652	752	1,068
per period in millions of $			
Market Structure			
Markets with no plants	2%	5%	1%
Markets with 1 plant	71%	64%	42%
Market with 2 plants	19%	20%	32%
Markets with more than 2 plants	8%	12%	26%

Table 12: Static Effect of Demand Smoothing

Most dramatically, the share of large plants in the industry increases from 31% to 47% when the 5 year smoothing policy is implemented, while the share of medium plants goes down from 16% to 5% and the share of small plants stays about the same. These larger plants have greater capacities are are more likely to be found in larger markets. Furthermore the number of plants increases from 2,700 to 4,339, inducing an increase in fixed cost expenditures from $ 652 million to $ 1,068 million per year.

The first effect, the presence of more large plants when either the 5 year or the constant demand policies are implemented, is due to the fact that demand smoothing policies can alter market size by changing the net present value of demand in each market, and this change in the effective market size alters the firm size distribution. For instance, in the constant demand counterfactual, markets that currently have high demand will have a much higher net present value of demand since they retain their high demand level forever. Likewise, markets that have low demand will retain this low level demand in perpetuity. This causes the number of plants per market is more dispersed under constant demand, with 5% of markets having no plants, versus 2% in the base case, and 12% of markets having more than 2 plants, versus 8% in the base case. While in principle I could find either more small plants or more big plants given the distribution of demand in different market, the net effect is to allow more large plants to be sustained.

The second effect is the "market expansion" effect of demand smoothing. When I regress log plants in a market on log construction employment, I find a coefficient of 0.69, indicating that a 1% increase in construction employment leads to less than a 1% increase in the number of firms in the market. This concavity of the relationship between the number of plants in a market and construction demand, leads to the implication that smoothing demand will lead to more plants in the market. I speculate that this effect is due to congestion costs for concrete deliveries when demand is particularly high. Congestion could be due to greater costs to make large number of deliveries at the same time because of labor and machinery shortages, or because some deliveries cannot be made during the weeks of the year when demand peaks. When yearly demand is higher, it is more likely for congestion to occur. Thus lower demand volatility leads to higher profits in the industry (holding market structure fixed), inducing a "market expansion" effect of reduced demand volatility. [36]

This increase in the number of plants changes market structure: the number of markets served by a more than one plant rises from 27% to 58%. Decreasing the number of monopoly markets will also have a positive effect on consumer surplus in the ready-mix concrete market.

7.3 Consumer and Producer Surplus

For the 29% of markets which were formerly monopoly markets, but became markets subject to competition, consumers of concrete would pay about 4% less on concrete based on the estimates in Figure 2. This would transfer 64 million dollars from producers of concrete to consumers per year holding

[36]If I allocate construction employment evenly over several years I find a higher net present value of log construction employment. Note that the fact that I use log construction demand rather than unlogged construction demand accentuates this effect. However, I find that log construction demand is a much predictor of the number of firms in a market than unlogged construction demand in the small markets used in this paper.

purchases of concrete fixed. This is a lower bound on the change in consumer surplus since any elasticity in the demand for concrete would further increase consumer surplus. Thus consumers of concrete would benefit from a reduction in demand volatility.

I also look at the effect reducing demand fluctuations on producer surplus. Note that ex-ante the effect on producer surplus is quite ambiguous since in this model of oligopoly dynamics with entry, producer surplus is generated by 1- the fact that incumbents have already paid sunk entry costs, thus earning profits that potential entrants cannot replicate, and 2- the integer constraint on entry, which allows a monopolist to earn profits, without the market being profitable enough to support a duopoly. Producer surplus is just the net present value of profits both for potential entrants and for incumbents. [37]

I find that producer surplus for incumbents would increase from 3.3 billon dollars in current net present value terms without the demand smoothing policy, to 4.0 billion dollars in NPV with the demand smoothing policy, representing an addition of 19%. I also compute producer surplus for potential entrants, who represent 80% of the "firms" in the data, which increases from

[37] To compute producer surplus I reformulate the problem in terms of choice specific value functions. Thus producer surplus is just:

$$PS = \sum_{i \text{ is incumbent}} V^i(x^{i0}) + \sum_{t=0}^{\infty} \beta^t \sum_{i \text{ is entrant}} V^i(x^{it}) \tag{21}$$

which is just the ex-ante value function for incumbents, plus the discounted value of entrants in the future which needs to be kept track of since I assume that if an entrant does not enter they get continuation value of 0. The ex-ante value function is $V(x) = \sum_{j \in A_i} W(j|x)\Psi(j|x) + \gamma - \sum_{j \in A_i} \ln(\Psi(j|s))\,\Psi(j|s)$.

163 billion dollars in the world with fluctuations to 172 billion dollars when the demand smoothing policy is used, or a 5% increase in surplus. Yet the surplus numbers for potential entrants are suspect, since the vast majority of this surplus is derived from 98.7% of potential entrants who choose never to enter, yet receive a payoff from their private information shock ϵ_{a_0}. Surplus from firms that do not enter is truly an artifact of the model, since how do we interpret the profits of firms that don't exist and choose never to enter. [38]

I also look at the level of investment that the industry can support. Table 11 shows that demand smoothing would raise annual sunk cost expenditures on entry and and size changes from 300 to 518 million dollars per year and raise fixed costs expenditures from 652 million per year to 1,068 million. The combined effect of sunk and fixed costs expenditures raises investment by 12%. Again this increase in expenditures is almost entirely due to the market expansion effect of the 5 year demand smoothing policy rather than the effect on the size distribution of the industry.

[38] Potential entrants represent more than 80% of the players in the game, and potential entrants who choose to enter are 1.3% of all potential entrants as is illustrated in the first row of Table 5 on page 29. Note that these figures really crucially on my assumption that there are 10 firms in each market, so the number of potential entrants are given mechanically as 10 minus the number of incumbents.

8 Conclusion

Fluctuations in demand for ready-mix concrete have a large effects on the composition, size and investment level in this industry. I considered a policy by which the government would sequence its construction budgets in such a way as to eliminate 5 year changes in demand, but the industry would keep long-run movements in demand.

To look at the effect of this policy of reducing short-run changes in demand, I estimated a oligopoly model of entry/exit and discrete investment. These estimates showed very large sunk costs, both for changing plant size and for exiting the market. Demand has a greater effect on large plants than on small plants, but there are smaller costs of building a small plant than a large plant. As market size grows, both the number of plants and the share of large plants increases. Finally there is a large effect of competitors, and thus the number of firms is pinned down by the demand and the effect of competition.

I find that this demand smoothing policy would have a large effect. The number of plants would increase by 47%. The share of large plants in the industry would increase from 31% to 47% while the share of medium plants would fall from 16% to 5%. This increase in both the number and size of plants is responsible for increasing both investment and fixed costs expenditures by 47%, from $ 0.95 billion to $ 1.6 billion per year. As well, there would be an increase in producer surplus among incumbents of 19%, raising it from

3.3 billion to 4.0 in discounted net present value. Moreover, 29% of markets change from monopoly markets to competitive markets, which would reduce consumer expenditures on ready-mix concrete by 64 million dollars per year holding quantity purchased fixed. I find that the number of plants in a market is a concave function with respect to demand. Thus a reduction in demand volatility induces a "market expansion" effect. This market expansion effect is similar to an increase in market size, raising the number of plants and the fraction of plants which are large.

Surprisingly, there would be almost no change in the dynamics of the ready-mix concrete industry, i.e. the rate at which firms change their size and shut down plants. The reason is that sunk entry and size changing costs are very large for this industry, this is it costly to react to demand shocks. Furthermore, firms are unlikely to react to demand shocks when demand is very volatile, since these demand shocks covey little information on the net present value of demand looking into the the future. As the process for demand becomes more predictable, firms increase their sensitivity to demand shocks. Thus the reduction in the magnitude of demand shocks, which would reduce turnover, is counteracted by an increase in the sensitivity of firms to current demand.

This paper showed that there are significant effects associated with even short-run demand volatility. For the ready-mix concrete industry one would see an industry which is 47% larger in terms of either investment or number of firms if a demand smoothing policy were implemented. Especially in a context where firms have high sunk costs, we might expect the consequences of volatility to be expressed not in higher turnover or more volatile investment, but in the level of investment itself. Finally, the oligopoly structure of the markets is key to evaluate the response of the industry, since competition in the product market determines the number of firms that can survive and the type of firms which are selected into the industry.

References

Aguiregabiria, V., and P. Mira (2007): "Sequential Estimation of Dynamic Games," *Econometrica*, 75(1), 1–54.

Bajari, P., L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

Bertsekas, D., and J. Tsitsiklis (1996): *Neuro-Dynamic Programming.* Athena Scientific.

Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite (2010): "Learning-By-Doing, Organizational Forgetting, and Industry Dynamics," *Econometrica*, 78(2), 453–508.

Bresnahan, T. F., and P. C. Reiss (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), 33.

Campbell, J., and H. Hopenhayn (2005): "Market Size Matters," *The Journal of Industrial Economics*, 53(1), 1–25.

Collard-Wexler, A. (2009): "Productivity Dispersion and Plant Selection in the Ready-Mix Concrete Industry," *Working Paper, New York University.*

Doraszelski, U., and M. Satterthwaite (2010): "Computable Markov-Perfect Industry Dynamics," *RAND Journal of Economics*, 41(2), 215–243.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Xu (2006): "Entry and Exit in Geographic Markets," Working Paper, Penn State University.

Dunne, T., M. J. Roberts, and L. Samuelson (1988): "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, 19(4), 495–515.

Ericson, R., and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, 62(1), 53–82.

Fershtman, C., and A. Pakes (2004): "Finite State Dynamic Games with Asymmetric Information: A Computational Framework," .

Gourieroux, C., and A. Monfort (1996): *Simulation-based econometric methods*. Oxford University Press, USA.

Gourieroux, C., and A. Montfort (1993): "Indirect inference," *Journal of Applied Econometrics*, 8, S85–S118.

Gowrisankaran, G. (1999): "Efficient Representation of State Spaces for Some Dynamic Models," *Journal of Economic Dynamics and Control*, 23(8), 1077–1098.

Griliches, Z., and J. Mairesse (1998): "Production Functions: The Search for Identification," *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, pp. 169–203.

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60(3), 497–529.

Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *The Review of Economic Studies*, 61(2), 265–289.

Jarmin, R. S., and J. Miranda (2002): "The Longitudinal Business Database," Working Paper, Center for Economic Studies, US Census Bureau.

Jovanovic, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50(3), 649–670.

Keane, M., and A. Smith Jr (2003): "Generalized Indirect Inference for Discrete Choice Models," *Working Paper, Yale University.*

KOSMATKA, S. H., B. KERKHOFF, AND W. C. PANARESE (2002): *Design and control of concrete mixtures*, Engineering bulletin ; 001. Portland Cement Association, Skokie, Ill., 14th edn.

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of the Multinomial Probit Without Numerical Integration," *Econometrica*, 57(5), 995–1026.

PAKES, A., S. BERRY, AND M. OSTROVSKY (2007): "Simple estimators for the parameters of discrete dynamic games, with entry/exit examples,," *RAND Journal of Economics*, 38(2), 373–399.

PAKES, A., AND R. ERICSON (1998): "Empirical Applications of Alternative Models of Firm and Industry Dynamics," *Journal of Economic Theory*, 79(1), 1–45.

PAKES, A., AND P. MCGUIRE (2001): "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'Curse' of Dimensionality," *Econometrica*, 69(5), 1261–1281.

PESENDORFER, M., AND P. SCHMIDT-DENGLER (2008): "Asymptotic Least Squares Estimators for Dynamic Games," *The Review of Economic Studies*, 75, 901–928.

POWELL, W. B. (2007): *Approximate Dynamic Programming: Solving the Curse of Dimensionality*, Wiley Series in Probability and Statistics. John Wiley and Sons.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

RYAN, S. P. (2006): "The Costs of Environmental Regulation in a Concentrated Industry," *Working Paper, M.I.T.*

SEIM, K. (2006): "An empirical model of firm entry with endogenous product-type choices," *The RAND Journal of Economics*, 37(3), 619–640.

SYVERSON, C. (2004a): "Market Structure and Productivity: A Concrete Example," *Journal of Political Economy*, 112(6), 1181–1222.

——— (2004b): "Product substitutability and productivity dispersion," *Review of Economics and Statistics*, 86(2), 534–550.

A Market Fixed Effects

In the main model, I use a market categories model which is meant to to mimic the inclusion of market fixed effects. These market fixed effects are critical to the estimation of the model since persistent market level differences in profitability lead to upward bias on the effect of competition. This bias, especially when it induces positive effects of competition, leads to very aberrant industry dynamics such as having a market flip between 0 and 10 plants due to a positive externality due to competition. The goal of this section is to motivate the use of market category effects based on the average number of firms in a market over time, and explain why other plausible corrections for market fixed effects using average construction employment or the number of plants in a pre-period, do not give the right answer.

I consider the following different specifications of the market category effects:

a) **No Market Effects**.

b) **Average Number of Firms in Market (rounded to nearest integer).** In the main estimates of the model, I use the average number of firms in the market rounded to the nearest integer. However, this approach suffers from an endogeneity problem. To put it most clearly, consider the following dynamic, two firm model of the type:

$$a_{it} = \alpha a_{-it} + \beta a_{it-1} + \epsilon_{it} \tag{22}$$

If I include a_{it+1} in the above regression, then I am including an endogenous regressor since a_{it+1} is a function of a_{it} which in turn depends on ϵ_{it}, and more broadly on the entire history of $\epsilon_{i\tilde{t}}$ for $\tilde{t} < t$.

c) **Average Number of Firms in Market in years before this one (rounded to nearest integer).** However, if I include lagged a_{it-2} then this is not an endogenous regressor, since there is no dependance on a_{it-2} except through a_{it-1} which is already included in the regression. The only issue with using the average number of firms in the market in previous years is that a market can switch categories over time which makes for a more difficult state space to deal with, which is the reason that I do not use these market category controls in the main part of the paper.

d) **Average Number of Firms in the 1976-1983 period, with data on the post 1983 period**. Notice that for this model, I am using the early period to condition the number of firms in the market. This is a version of model c), but but the pre-period on which I condition does not change within a market.

e) **Average Construction Employment**. Here I classify markets by the average level of construction employment from 1976-1999. This is an exogenous classification scheme since it does not depend on what ready-mix concrete firms are doing.

f) **Market Fixed Effects (Conditional Logit).**

Table 13 presents estimates from the binary logit model of entry and exit for specifications (a)-(f). I have chosen the binary logit model since it allows me to use the conditional logit with market fixed effects. [39] Column (a) shows estimate without market category controls (henceforth referred to as no market effects), while column (f) shows estimates with market fixed effects (henceforth referred to as market fixed effects), while columns (b)-(e) show different market category controls. Columns (b) and (c), i.e. with market controls based on the average number of plants and the average number of plants in the periods before this one, are similar to the market fixed effect estimates in column (f). Likewise, columns (d) and (e) show estimates that are more similar to the no market effects estimates in column (a).

The effect of past plant size on activity are fairly similar in all of these estimates, with smaller effects of plant size in the market fixed effect specifications (f), (b) and (c) than the no market effect specification (a), (d) and (e). Unobserved heterogeneity between markets is loaded onto variable indicating state dependence, such as past plant size. The effect of log construction employment is higher at 0.133 to 0.099 in the no market effect models (a), (d) and (e) than in the market fixed effect estimates, which have estimates from 0.033 to −0.034. These higher effects of demand are due to the fact that firms are far more likely to react to cross-sectional differences in demand (which are more likely to be persistent) than to year to year changes in demand. Likewise, the effect of the second competitor (which will representative of the effect of competition more broadly) varies from −0.074 to 0.003 in the no market effect columns (a), (d) and (e), but ranges from −0.635 to −0.529 in the market fixed effect columns (f), (b) and (c). This is indicative of the fact that unobserved differences in the profitability of a market will be correlated with the number of plants in the market.

There are two main conclusions from the table that are relevant for my choice of market categories. First, the market categories based on the either the average number of firms (b) or the average number of firms in all periods before today (c) do a good job in mimicking true market fixed effects. However, using categories based on the number of firms before 1983 (d), or using information about the average level

[39] Technically, I can also use a multinomial conditional logit, but the number of categories I need to condition on becomes fairly large. As well, I am not presenting marginal effects here since the conditional logit does not estimate the market fixed effects.

of construction demand (e) do not replicate market fixed effect estimates, and in fact mimic no having any market controls whatsoever. Second, while it is true that using the average number of firms over time conditions on an endogenous variable, I can equally easily use the lagged number of firms which does not condition on an endogenous variable and obtain virtually identical results. Thus the issue of endogeneity is of limited practical importance in the use of the average number of firms over time as a grouping.

Dependant Variable: Activity	(a)	(b)	(c)	(d)	(e)	(f) Conditional Logit
Log County	0.133***	-0.034**	-0.034***	0.129***	0.099***	0.033
Construction Employment	(0.011)	(0.011)	(0.010)	(0.015)	(0.019)	(0.023)
First Competitor	-1.403***	-1.805***	-1.748***	-1.306***	-1.421***	-2.002***
	(0.052)	(0.051)	(0.048)	(0.066)	(0.052)	(0.043)
Second Competitor	0.003	-0.529***	-0.553***	-0.074	-0.008	-0.635***
	(0.036)	(0.037)	(0.036)	(0.047)	(0.037)	(0.030)
Third Competitor	0.026	-0.359***	-0.384***	-0.071	0.027	-0.394***
	(0.044)	(0.044)	(0.043)	(0.058)	(0.044)	(0.036)
Log Competitors above 4	0.022	-0.118***	-0.170***	-0.001	0.035	-0.187***
	(0.029)	(0.028)	(0.028)	(0.040)	(0.029)	(0.025)
Small	5.889***	5.703***	5.720***	5.977***	5.887***	5.585***
	(0.037)	(0.035)	(0.035)	(0.047)	(0.037)	(0.025)
Small,	5.665***	5.388***	5.393***	5.707***	5.657***	5.220***
Medium in Past	(0.048)	(0.045)	(0.045)	(0.057)	(0.048)	(0.033)
Small,	4.866***	4.636***	4.643***	4.944***	4.865***	4.450***
Large in Past	(0.065)	(0.063)	(0.062)	(0.075)	(0.065)	(0.041)
Medium	7.503***	7.292***	7.315***	7.696***	7.495***	7.234***
	(0.057)	(0.055)	(0.055)	(0.075)	(0.057)	(0.050)
Medium,	7.511***	7.237***	7.251***	7.585***	7.503***	7.122***
Large in Past	(0.080)	(0.079)	(0.079)	(0.094)	(0.081)	(0.074)
Large	7.671***	7.446***	7.450***	7.724***	7.676***	7.436***
	(0.056)	(0.054)	(0.054)	(0.068)	(0.056)	(0.050)
Market						
Classification Variable						
Average Number of Plants		X				
Lagged Average Plants			X			
Before 1983 Average Plants				X		
Construction Employment					X	
Category 2		1.053***	1.118***	0.225***	0.132**	
		(0.036)	(0.032)	(0.062)	(0.049)	
Category 3		1.668***	1.836***	0.348***	0.199**	
		(0.050)	(0.047)	(0.058)	(0.061)	
Category 4		2.293***	2.424***	0.482***	0.169*	
		(0.063)	(0.062)	(0.061)	(0.082)	
Constant	-3.805***	-2.985***	-2.970***	-4.089***	-3.715***	
	(0.065)	(0.066)	(0.062)	(0.089)	(0.090)	
Observations	409850	409850	409850	260170	409850	409850
Markets	2029	2029	2029	2014	2029	2029
Log-Likelihood	-45695	-44483	-44304	-27334	-45682	-39670
χ^2	44067	47153	46207	29860	44985	284475

(Standard Errors are Clustered by Market).

Table 13: Market Effects in the Binomial Logit Regression of Entry and Exit

B Discrete Action Stochastic Algorithm: Termination Criteria

The stopping rule is based on the fact that if I have the "correct" W function, then it will satisfy the Bellman equation. However, it is computationally expensive to calculate the W-function exactly, instead we can approximate the value function using forward simulation. Consider the locations $R \subset S \times A$ defined as the state-action pairs visited in the last 1 million iterations (keep a hit counter that tracks the last 1 million iterations denoted $rh(l)$).

Algorithm Fershtman-Pakes Stopping Rule (FPStop)

For all locations $l = \{a_i, x\} \in L$ which have been visited in the last 1 million iterations:

1. Compute the W-function using a one step forward simulation. For $k = 1, ..., K$ (I use $K = 10\,000$):
   (a) Draw an action profile a^k and a state tomorrow $x^{k\prime}$ given location l.
   (b) Get rewards:
   $$\begin{aligned} R^k =& r(a^k|x^{k\prime}, \theta) + \tau(a_i^k|x_i, \theta) \\ &+ \beta \sum_{j \in A} W(j, x^{k\prime}) P[j|x^{k\prime}] \\ &+ \beta \left(\gamma - \sum_{j \in A} \ln(P[j|x^{k\prime}]) P[j|x^{k\prime}] \right) \end{aligned} \tag{23}$$
   (c) Compute the approximation to the W-function:
   $$\tilde{W}(l) = \frac{1}{K} \sum_{k=1}^{K} R^k \tag{24}$$
2. Compute the difference in value functions weighted by the recent hit counter rh:
   $$\gamma = \frac{1}{\sum_l rh(l)} \sum_l rh(l) * (\tilde{W}(l) - W(l))^2 \tag{25}$$

If the test statistic γ is small enough, then we can argue that we have a good approximation. In practice I have used the fact that the recent hit counter weighted

R^2 between $\tilde{W}(l)$ and $W(l)$ is greater than 0.999. This usually happens after as little as 50 million iterations, and it is usually more efficient to run the DASA for 150 million iterations (i.e. 15 minutes) which will lead to a W function which satisfy the FPStop criteria. Furthermore, in this application there are only about 3 000 state-action pairs (where the action is not 0) that are visited in the last 1 million iterations. Thus the ergodic class R is quite small compared to the size of the entire state space.

C Modified DASA to Compute the Gamma function

I use a modified DASA to compute the Γ function. The two differences are that (i) I shut down the policy function update in the DASA, and (ii) I compute the net present value of the components of rewards rather than the rewards themselves (which would require me to have information on the parameters θ).

Algorithm Γ-Compute Discrete Action Stochastic Algorithm (**GC-DASA**)

1. Start in a location $l_0 = \{a_0, x_0\}$.

2. Draw an action profile $a|a_i \sim 1(a_i = a_0)\prod_{-i}\hat{P}[a_{-i}|x]$ and a state in the next period x' given action profile a:

$$x'|a \sim \hat{D}[M'|M]\prod_i \iota(x'_i|a_i, x_i) \tag{26}$$

   where $\iota(x'_i|a_i, x_i)$ is the *updating* function, which updates the firm's state based on a firm's action and the firms largest size in the past.

3. Increment the hit counter (how often you have visited the state-action pair): $h(l) = h(l) + 1$.

4. Compute i^{th} component of payoffs R^i of the action a_i as:

$$\begin{aligned} R^i =& r^i(a_i, x) \\ &+ \beta\sum_{j\in A}\Gamma^i(j, x')P[j|x'] \end{aligned} \tag{27}$$

5. Update the Γ-function:

$$\Gamma^{i\prime}(l) = \alpha R^i + (1-\alpha)\Gamma^i(l) \tag{28}$$

where $\alpha = \frac{1}{h(l)}$.

6. Update current location to $l' = \{a'_i, x'\}$.

7. The stopping rule is Fershtman and Pakes (2004)'s.

D Simulated Indirect Inference Estimation

The simulated indirect inference estimator used in equation (35) on page 76 uses the choice probabilities $\Psi(a|x,\Gamma,\theta)$ as an outcome vector, i.e. $\tilde{\mathbf{y}}_{\mathbf{n}} = \Psi(a|x,\Gamma,\theta)$. Typically, one would sample outcomes $\mathbf{y_n}$ from the choice probabilities $\Psi(a|x,\Gamma,\theta)$. I can show that using the $\tilde{\mathbf{y_n}}$ is equivalent to sampling actions as the number of actions tends to infinity.

Denote the outcome vectors $\mathbf{y_n^s}$ as:

$$\mathbf{y_n^s} = \begin{bmatrix} 1(a_n^s = \text{small}) \\ 1(a_n^s = \text{medium}) \\ 1(a_n^s = \text{big}) \end{bmatrix} \tag{29}$$

where the action $a_n^s \sim \Psi(\cdot|x,\Gamma,\theta)$ is drawn from the choice probabilities Ψ. The simulation draws are indexed from $s = 1, \cdots, S$. The $\beta^S(\theta)$ coefficient is estimated using outcome vectors $\{y_n^s\}_{\{s=1,\cdots,S\},n}$. The criterion function using S simulation draws of actions is thus:

$$\mathcal{Q}^S(\theta) = \left(\hat{\beta} - \tilde{\beta}(\theta)\right)' \mathbf{W} \left(\hat{\beta} - \tilde{\beta}(\theta)\right) \tag{30}$$

D.1 Consistency Proof

In this section I will show conditions under which the procedure I use in this paper is a consistent estimator of θ. Specifically, I will show the conditions that need to be satisfied for Proposition 1 on page S89 in Gourieroux and Montfort (1993) dealing with the consistency of indirect inference estimators, to be satisfied.

Define the criterion function used to compute $\tilde{\beta(\theta)}$ (for a given value of θ) as:

$$\begin{aligned}\mathcal{S}^{N,K}(\beta,\theta) = &\sum_{n=1}^{N}\sum_{k=1}^{K}\left[1\left(a_n^k = \text{small}\right) - Z_n\beta_s\right]^2 \\ &+ \left[1\left(a_n^k = \text{medium}\right) - Z_n\beta_m\right]^2 + \left[1\left(a_n^k = \text{large}\right) - Z_n\beta_l\right]^2\end{aligned} \tag{31}$$

where N denotes the number of observations and K denotes the number of simulation draws to draw actions a_n^k from the policy function $\psi(a_n|x_n,\theta,\Gamma(\hat{P}^N,\hat{D}^N))$.

Note that $\mathcal{S}^{N,K}(\beta,\theta)$ is the criterion used in OLS estimation, just the sum of squared errors.

The first step is to show that I can replace draws of a_n^k with the actual policy function ψ, or in other words $\mathcal{S}^{N,K}(\beta,\theta) \rightarrow^{\text{a.s.}} \mathcal{S}^{N,\infty}(\beta,\theta)$ uniformly as $K \rightarrow \infty$.

Theorem 4 *As the number of simulation draws K tends to infinity, $\mathcal{S}^{N,K}(\beta,\theta) \rightarrow^{a.s.} \mathcal{S}^{N,\infty}(\beta,\theta)$ uniformly.*

Proof: I will show the proof using only the choice to be small to lighten the notation, but the proof extends to as many actions as I want:

$$\begin{aligned}\mathcal{S}^{N,K}(\beta,\theta) =& \sum_{n=1}^{N}\frac{1}{K}\sum_{k=1}^{K}\left[1\left(a_n^k = \text{small}\right) - Z_n\beta_s\right]^2 \\ =& \sum_{n=1}^{N}(Z_n\beta_s)^2 + \sum_{n=1}^{N}\sum_{k=1}^{K}\frac{1}{K}1\left(a_n^k = \text{small}\right)^2 - 2\sum_{n=1}^{N}Z_n\beta_s\sum_{k=1}^{K}\frac{1}{K}1\left(a_n^k = \text{small}\right)\end{aligned} \tag{32}$$

As $K \rightarrow \infty$, $\sum_{k=1}^{K}\frac{1}{K}1\left(a_n^k = \text{small}\right) \rightarrow \psi(a_n = \text{small}|x_n,\theta,\Gamma(\hat{P},\hat{D}))$ since this is just an average, and $\sum_{k=1}^{K}\frac{1}{K}1\left(a_n^k = \text{small}\right)^2 \rightarrow \psi(a_n = \text{small}|x_n,\theta,\Gamma(\hat{P},\hat{D}))^2$. Fix me Thus I can rewrite $\mathcal{S}^{N,\infty}(\beta,\theta)$ as:

$$\begin{aligned}\mathcal{S}^{N,\infty}(\beta,\theta) =& \sum_{n=1}^{N}(Z_n\beta_s)^2 + \sum_{n=1}^{N}\psi(a_n = \text{small}|x_n,\theta,\Gamma(\hat{P},\hat{D})^2 \\ & - 2\sum_{n=1}^{N}Z_n\beta_s\psi(a_n = \text{small}|x_n,\theta,\Gamma(\hat{P},\hat{D})) \\ =& \sum_{n=1}^{N}\left[\psi(a_n = \text{small}|x_n,\theta,\Gamma(\hat{P},\hat{D})) - Z_n\beta_s\right]^2\end{aligned} \tag{33}$$

Second, I need to show that $\mathcal{S}^{N,\infty}(\beta,\theta) \rightarrow^{\text{a.s.}} \mathcal{S}^{0,\infty}(\beta,\theta)$ as $N \rightarrow \infty$. The first condition is that the linear probability estimator is consistent, which is just an outcome of the OLS estimator being a consistent estimator, which is a standard proof. However, I am not using the true $\Gamma^0(P^0,D^0)$ but an estimate of $\Gamma(\hat{P},\hat{D})$ due to sampling error in the conditional choice probabilities P and the demand transition process D, as well as approximation error in the computation of Γ. The CCP's $\hat{P}^N \rightarrow P^0$ which happens since I am using a consistent estimator of the CCP's, just a parametric multinomial logit, which is consistent using the usual proofs on the consistency of M-estimators. Likewise $\hat{D}^N \rightarrow D^0$ as $N \rightarrow \infty$ since I am using a consistent estimator of D, just a bin estimator where the number of bins is fixed as $N \to \infty$. Now the next point is to show that $\Gamma^L(P^0, D^0) \to \Gamma^0(P^0, D^0)$ as the number of iterations L in the DASA goes to infinity. It will be difficult to show convergence of the DASA, since to my knowledge there is no proof of the convergence of algorithms that compute the solutions to games (in contrast to single agent problems). However, the Fershtman and Pakes (2004) convergence criterion can be used to check the convergence of the DASA, and I can send the tolerance of the Fershtman-Pakes criterion to 0 as $N \to \infty$. [40]

The convergence of $\Gamma^L(P^0, D^0) \to \Gamma(P^0, D^0)$ implies the convergence of $\mathcal{S}^{N,K}(\beta, \theta) \to \mathcal{S}^{\infty,\infty}(\beta, \theta)$ as $K \to \infty$ and $N \to \infty$. This satisfies assumption (A2) in Indirect Inference.

Assumption (A3) of Indirect Inference requires that:

$$\tilde{\beta}(\theta) = \text{argmax}_\beta \mathcal{S}^{\infty,\infty}(\beta, \theta) \tag{34}$$

be a continuous function and have a unique value. Continuity is an outcome of the OLS structure of $\mathcal{S}$, while uniqueness occurs if Z_n is full rank and the dimension of β is smaller than the dimension of Z_n.

The final condition, (A4) requires that $\tilde{\beta}(\theta)$ be one to one and have full rank . I will assume this condition, but notice that the dimension of β is larger than the dimension of θ and I have checked that $\tilde{\beta}(\theta)$ is full rank in the estimation of model.

Since conditions (A1), (A2), (A3) and (A4) are satisfied, then $\hat{\theta}$ defined as the minimizer of:

$$\mathcal{Q}(\theta) = \left(\hat{\beta} - \tilde{\beta}(\theta))\right)' \mathbf{W} \left(\hat{\beta} - \tilde{\beta}(\theta)\right) \tag{35}$$

will be a consistent estimator of θ as $N \to \infty$.

[40] Notice that since there is a full support shock ϵ to the payoffs of any actions, Γ is computed correctly on the entire state space S, since the set of recurrent points is the entire state space, i.e. $S = R$. The DASA used to compute Γ is a version of the Q-learning algorithm, where consistency proof are provided for the single agent (non-game version) in Propositions 5.5 and 5.6 on page 248-249 in Bertsekas and Tsitsiklis (1996) show conditions under which the DASA's (which is the game version of a Q-learning algorithm) computation of $\Gamma^{\cdot}$ converges with probability one to Γ^0. These conditions are (1) that policies are proper, i.e. there is a positive probability that a firm will exit after t period, which is true in this context due to the full support of the shock distribution for each action, including the choice to exit; and (2) for improper policies, there is a negative infinite value of W for at least one state. Unfortunately, there is to my knowledge no proof with shows the convergence of the Q-learning algorithm in the context of a game.

E Price Data

The Census Bureau does not generally collect price data. This job is left to the Bureau of Economic Analysis and the Bureau of Labor Statistics. However, following Syverson (2004a) we can generate prices using the following equation:

$$p_{it}(c) = \frac{s_{it}(c)}{q_{it}(c)} \tag{36}$$

which is just sales of the commodity divided by quantity sold. While these "prices" may be good indicators of price dispersion (the application Syverson considers), they are particular poor measures of actual plant prices, with an interquartile range over 2 log points (the third quartile is 100 times bigger than the first price quartile). This is probably because of how measurement error in the numerator and especially the denominator interact.

To reduce the impact of imputed data and measurement error on the dispersion of prices, I apply a version of Syverson (2004b)'s procedures:

1. Hot Imputes in the data are identified as prices that satisfy the following:

   $$|p_i^t - p_j^t| < 0.0001 \text{ for some } i \text{ and } j \text{ in the data} \tag{37}$$

   I drop all prices that are hot imputes. Notice that this procedure will also eliminate cold imputes, defined as prices which equal the mode in the current year.

2. I trim the data by dropping observations that are less than 1/5 or more than 5 times the median price for the current year.

The deflated data is computed by $p_i^{Dt} = p_i^t / cpi^t$ where I normalize the cpi in 1977 to be equal to 1 (i.e. $cpi^t = \text{raw cpi}^t / \text{raw cpi}^{1977}$). This eliminate differences in price level across time, but does not incorporate differences in prices between regions.

F Maximum Likelihood Estimates

In this section I present estimates of the structural model in Table 8 on page 46 using a maximum likelihood criterion. Table 14 shows estimates of the model, where column I and II use CCPs with market category controls $\hat{P}^{\mu}$ to compute the Γ function, while column III and IV use CCPs without market controls $\hat{P}$ to compute Γ.

		$\hat{P}^{\mu}$ (Market Effects)				$\hat{P}$ (No Market Effects)			
		I.		II.		III.		IV.	
		Coef.	S.E.*	Coef.	S.E.	Coef.	S.E.*	Coef.	S.E.
Fixed Cost	Small	-462	(3)	-361	(6)	-575	(3)	-465	(6)
	Medium	-779	(5)	-721	(8)	-955	(4)	-850	(7)
	Large	-886	(6)	-834	(9)	-917	(5)	-810	(7)
Log Construction	Small	-45	(1)	-52	(2)	43	(1)	18	(2)
Employment	Medium	81	(2)	76	(2)	33	(1)	9	(1)
	Large	167	(2)	162	(3)	30	(1)	1	(2)
1st Competitor	Small	-229	(3)	-658	(7)	-840	(6)	-848	(6)
	Medium	-220	(5)	-628	(8)	-405	(7)	-414	(7)
	Large	-159	(6)	-530	(9)	-334	(7)	-345	(7)
Log Competitors	Small	-81	(2)	-33	(2)	265	(2)	263	(3)
(above 1)	Medium	-29	(2)	24	(3)	201	(3)	203	(3)
	Large	-41	(3)	-7	(3)	157	(3)	161	(3)
Market 2 Effect				203	(4)			71	(4)
Market 3 Effect				285	(5)			119	(6)
Market 4 Effect				255	(7)			140	(7)
Transition Costs									
Out →Small		-1640	(7)	-1616	(8)	-1151	(8)	-1156	(8)
Out →Medium †		-2000	(10)	-2000	(12)	-2000	(12)	-2000	(12)
Out →Large		-1996	(10)	-2006	(11)	-2242	(13)	-2243	(13)
Small →Medium		-217	(7)	-224	(7)	-145	(3)	-146	(3)
Small, Past Medium →Medium		-515	(10)	-563	(11)	-395	(9)	-395	(9)
Small, Past Large →Medium		-182	(9)	-192	(10)	-49	(3)	-51	(3)
Small →Large		-788	(22)	-860	(24)	-372	(16)	-373	(16)
Small, Past Medium →Large		-346	(15)	-377	(16)	-91	(4)	-88	(4)
Small, Past Large →Large		-251	(8)	-262	(9)	-156	(4)	-161	(4)
Medium →Small		-147	(5)	-155	(5)	-65	(2)	-63	(2)
Medium, Past Large →Small		-198	(4)	-210	(5)	-163	(3)	-169	(3)
Medium →Large		-273	(7)	-290	(8)	-107	(3)	-104	(3)
Medium, Past Large →Large		-130	(3)	-138	(4)	-90	(3)	-95	(3)
Large →Small		-287	(7)	-316	(8)	-160	(3)	-155	(3)
Large →Medium		-52	(3)	-56	(3)	-72	(2)	-69	(2)
Standard Deviation of Shock		372		411		1151		1151	
Observations		409848		409848		409848		409848	
Wald		2523812		2575628		4750171		4756505	
Log-Likelihood		-152249		-149174		-187326		-187038	

∗: Standard Errors computed assuming no first-stage error in the $\hat{P}$ conditional choice probabilities.

Table 14: Maximum Likelihood Estimates of the Dynamic Model