Essays in Regional Industrial Organization

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Brian Michael Adams

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Thomas J. Holmes, Adviser

August, 2013

Acknowledgements

I thank my committee, Erzo Luttmer, Amil Petrin, and Joel Waldfogel, and especially my advisor, Thomas Holmes, for their advice and guidance. I also received helpful suggestions from Kyoo-il Kim and Elena Pastrino. I thank Imke Reimers for reading drafts and suggesting improvements. I thank my wife, Emily Adams, for her love, support, and proofreading.

Chapter 2 is a joint work with Kevin Williams. I am grateful to have him as a coauthor. We thank the Minnesota Supercomputering Institute for the use of their computational resources.

I thank Christian Hilber and Zhu Wang for their comments as conference discussants on earlier versions of Chapter 3. I also thank seminar participants at the Bureau of Labor Statistics, California State University - East Bay, Towson University, and the University of Akron for their suggestions.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The dissertation is a collection of three essays in regional industrial organization. Chapter 2 analyses zone pricing in retail drywall. Retailers post the exact same price in all stores in large regions, even though these price zone encompass markets of different size and competiveness. The chapter estimates what the costs of making finer price zones must be based on forgone profits of greater price discrimination.

Chapters 3 and 4 examine how motor vehicle parts suppliers react to changes in assembly plant location. Chapter 3 estimates a model of location choice in which forward-looking parts supplier decide where to enter and when to exit based on their expectations of the evolution of the industry. This model identifies what attracts suppliers and how they will react to changes in the alignment of assembly plants. Chapter 4 uses a natural experiment to find how a new assembly plant opening affects the number of supplier plants and jobs within various distances.

Chapter 2

Zone Pricing and Spatial Menu Costs: Estimates for Drywall

2.1 Introduction

In the empirical industrial organization literature, the standard approach models prices as being set at the *market level.* Competing firms in a particular market each set a specific price for that market, taking as given the prices set by competing firms. In practice, retail firms commonly set prices at the *zone level.* These pricing zones, although usually geographically contiguous, often combine distinct markets that may be hundreds of miles apart, and that differ in significant ways. A firm's pricing zone might include urban and rural markets, markets with different degrees of competition, and markets where input costs vary substantially. With these differences across markets, we might expect the firm to set individual prices in each market rather than a common price throughout the zone.

In this chapter, we develop an empirical analysis of zone pricing. Our application is the retail drywall industry. We explain a number of features of drywall retailing that make it an ideal industry for such a study, and further we are able to obtain a uniquely rich data set for this industry. We introduce the concept of "spatial menu costs," which we interpret as a type of friction that induces firms to set a constant price over multiple markets; that is, a zone price. Through use of empirical techniques to estimate demand and cost conditions, we evaluate the magnitude of the spatial menu costs needed to induce firms to adopt zone pricing. We also evaluate how the existence of spatial menu costs affects the competitive interaction of firms, and connect our results to the earlier theoretical literature on price discrimination between competing firms.

The spatial menu costs considered here relate in spirit to the concept of menu costs prominent in the macroeconomics literature. The macro literature documents how prices change infrequently over time, and this inter-temporal price rigidity has potentially significant implications for the macroeconomy. Here, the price rigidity is across space – a pricing zone. Many of the key issues from the macro literature apply in our context.

Retail drywall markets have several features that aid analysis of zone pricing. In the mainstream retail sector for drywall, competition is between a few large chains, all of which practice zone pricing. Drywall, also known as wallboard or gypsum board, is costly to transport, so costs vary geographically. The distribution centers for each major retail chain have known locations, making costs predictable. Some observed pricing zones span multiple, diverse markets, making advertising an unlikely reason for uniform pricing. Some pricing zones also include monopoly markets and markets with multiple stores from each chain. The variation of costs, competition, and demand variables within a pricing region allow us to uncover the role of spatial menu costs in forcing pricing zones.

Our work contributes to the literature on price discrimination, as spatial menu costs are an interesting impediment to price discrimination. Firms have chosen not to set completely uniform prices and so engage in limited price discrimination. We show that without spatial menu costs, drywall retailers would set a discriminatory price in each market. In the standard monopoly setting, limiting the firm to a zone system can only lower profits. However, Holmes (1989) and Corts (1998) show that in environments where firms compete, the effect of being able to price discriminate has an ambiguous effect on profit; impediments to price discrimination limit what a particular firm can do, but it also affects what its rivals are doing. In our analysis we not only determine how profits change when market level pricing is permitted, holding fixed what competitors do, we also model the competitive interactions in choosing zone level or market level pricing.

We estimate a structural demand and supply model to recover these spatial menu costs. Consumers select between all available drywall products at each nearby store from either chain. We find drywall to be a highly substitutable product, but overall industry elasticity is very low. We estimate the marginal cost of each product in each store by accounting for transportation costs from the warehouse. With these demand and costs estimates, we can estimate profits under the current pricing zones and compare them to pre-menu cost profits in counterfactual equilibria where one or both other the firms instead uses market level pricing. Multiple equilibria exist for a given zone configuration. We compare the current regime with a small adjustment in zones to yield unique counterfactual equilibria. Aggregating across markets, we find the spatial menu costs to be 2.2% of current revenues. For the 128 stores in the sample, this equates to roughly $4.6 million in addition profits for retail drywall annually. The spatial menu costs that would rationalize the chains' decision not to separate stores into their own pricing districts are substantial, though small enough that managerial effort costs (as in Zbaracki, Ritson, Levy, Dutta, and Bergen (2004)) are a likely explanation.

Previous work on zone pricing is relatively sparse in both the economics and marketing literature. Montgomery (1997) and Chintagunta, Dubé, and Singh (2003) are important exceptions. They both examine a supermarket chain that practices zone pricing and ask how profits and consumer surplus would change if the chain switched from zone pricing to store-by-store pricing. Marginal costs were assumed to be the same at all stores. Although this is an acceptable abstraction for supermarkets within a city, such an assumption cannot be maintained with drywall. We build upon these works by allowing costs to be specific to each store. Another important difference is that instead of analyzing what is happening to one firm in isolation, holding fixed the environment of the firm as it changes from zone to store-by-store, this chapter takes into account how switching regimes can affect the entire competitive interaction. We find large profit gains for moving to store level pricing in an environment in which competitor's prices are locked into their current level. Without accounting for the competitive interaction of firms, we find that menu costs are overstated by as much as 32% or over $2.1 million annually for the stores examined.

This chapter is organized as follows: In Section 2, we document zone pricing for the home improvement retail industry and describe the data used for this study. In Section 3, we introduce the supply and demand system. In Section 4, we present estimation results and in Section 5, we conduct analysis on alternative pricing regimes.

2.2 Data and Descriptive Evidence

We create two new data sets for products available at the two largest home improvement warehouse retailers in the United States. First, to document zone pricing for drywall, plywood, and interior paint, we obtain a cross section of prices at every Lowe's and Home Depot store in the United States. To estimate an empirical model of drywall demand and pricing, we construct a more detailed data set of prices, sales quantities, and characteristics for all drywall products offered at 128 stores in the Intermountain West.

Pricing Patterns in National Data

Drywall prices posted on company websites on 30 April 2013 reveal the use of zone pricing by both Home Depot and Lowe's. For both chains, prices on a given product vary considerably at a national level, but are exactly the same price at all of the chain's stores within contiguous areas. These areas are largely the same for different drywall products. We define a pricing zone as an area in which stores have exactly the same price on all products offered. We show that although most pricing zones are small, some contain a large number of stores and span a diverse set of markets.

Price levels for regular 5/8" x 4' x 8' gypsum board are mapped at every Lowe's and at every Home Depot in Figure 2.1. Each dot on the map represents a store, and its color represents its price level. Nationally this product has considerable price variation, even though within some areas there is no price variation. There are only 126 distinct prices for 1714 Lowe's stores in the United States. Surprisingly in some areas, such as the Carolina Piedmont, a sharp boundary separates two zones with very different prices. Elsewhere, such as the upper Midwest, prices are similar over a wide area.

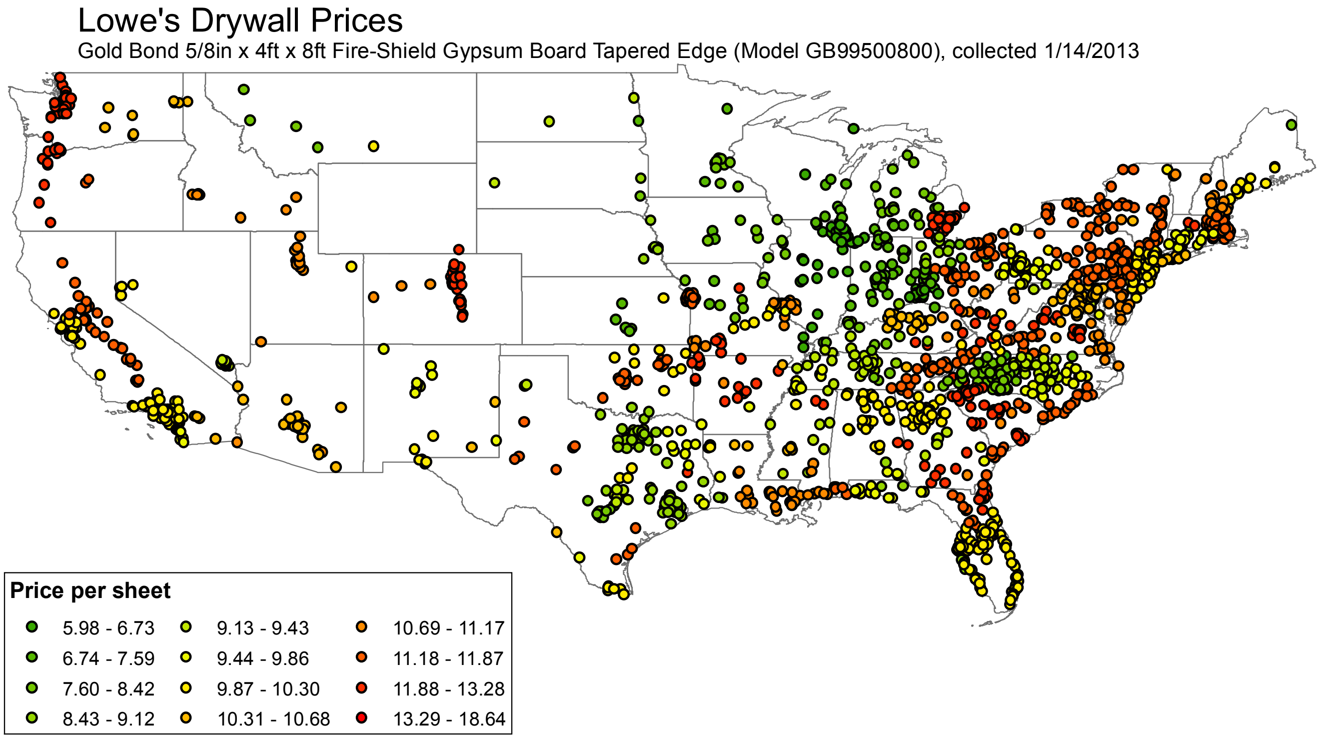
Figure 2.3 plots unique prices for a regular 5/8" x 4' x 8' gypsum board product available at Home Depot stores in the Western United States. Each dot again represents a store; stores with exactly the same price are the same color. The figure shows geographically contiguous pricing zones. For example, there is a unique price for this drywall product for all Home Depot stores in Oregon. All locations in Utah and Southern Idaho have the same price, although stores in eastern and western Washington have different prices. The unique pricing regions for different products within a chain are often but not always the same. The prices for 1/2” x 4’ x 12’ drywall board exhibit three prices in the state of Washington instead of two, but the two pricing zones in western Washington combine to correspond exactly into the pricing region for 5/8” x 4’ x 8’ drywall. Pricing zones are regions in which all drywall products have the same price. By definition, these pricing zones are no larger than the uniform price region for any product. Home Depot has 165 drywall pricing zones for 1,979 stores; Lowe’s has 129 drywall pricing zones for its 1,714 stores in the United States. Three pricing zones had the prices for regular 5/8” x 4’ x 8’ sheets that were seen in other zones, hence the 126 distinct prices for that product.

Many pricing zones contain only one metropolitan area, and these may be justified for marketing reasons or because costs and competition are similar within a city. This would lead to profit-maximizing prices being the same across stores. However, several pricing zones are much larger. Table 2.1 reports on the number of stores in drywall pricing zones by chain. The drywall pricing zones with the most stores for both Lowe’s and Home Depot are in Southern California, both extending from Los Angeles to San Luis Obispo. Lowe’s has 32 pricing zones that span more than 200 miles; Home Depot has 16 such zones. Pricing zones of such size represent multiple consumer markets. Drywall is bulky, making it unlikely that consumers would substitute to stores a great distance away.

Costs and market structure can also vary in vast pricing zones. Table 2.2 presents an example from a large drywall pricing zone based around Salt Lake City. The Home Depot stores in Logan, Utah, Rock Springs, Wyoming, and Elko, Nevada are all in this pricing zone, and hence, the prices for drywall within these stores are the same – the 5/8” x 4’ x 8’ drywall board is $10.98. The Home Depot in Logan faces competition from Lowe’s, located a mile away. The nearest Lowe’s to the stores in Rock Springs and Elko are 107 and 168 miles away. Further, at around 50 pounds per sheet of drywall, distance should play an important role in costs. The distance to the nearest distribution center and the distance to the nearest factory both vary by hundreds of miles. Profit maximizing prices for each of these stores should differ substantially, yet Home Depot places all three stores in the same zone and assigns identical prices.

The pricing zones for Lowe’s and Home Depot in the United States often have the same boundaries. Figure 2.2 plots unique prices for the same drywall product at Lowe’s

Figure 2.1: Maps of US Lowe's and Home Depot prices



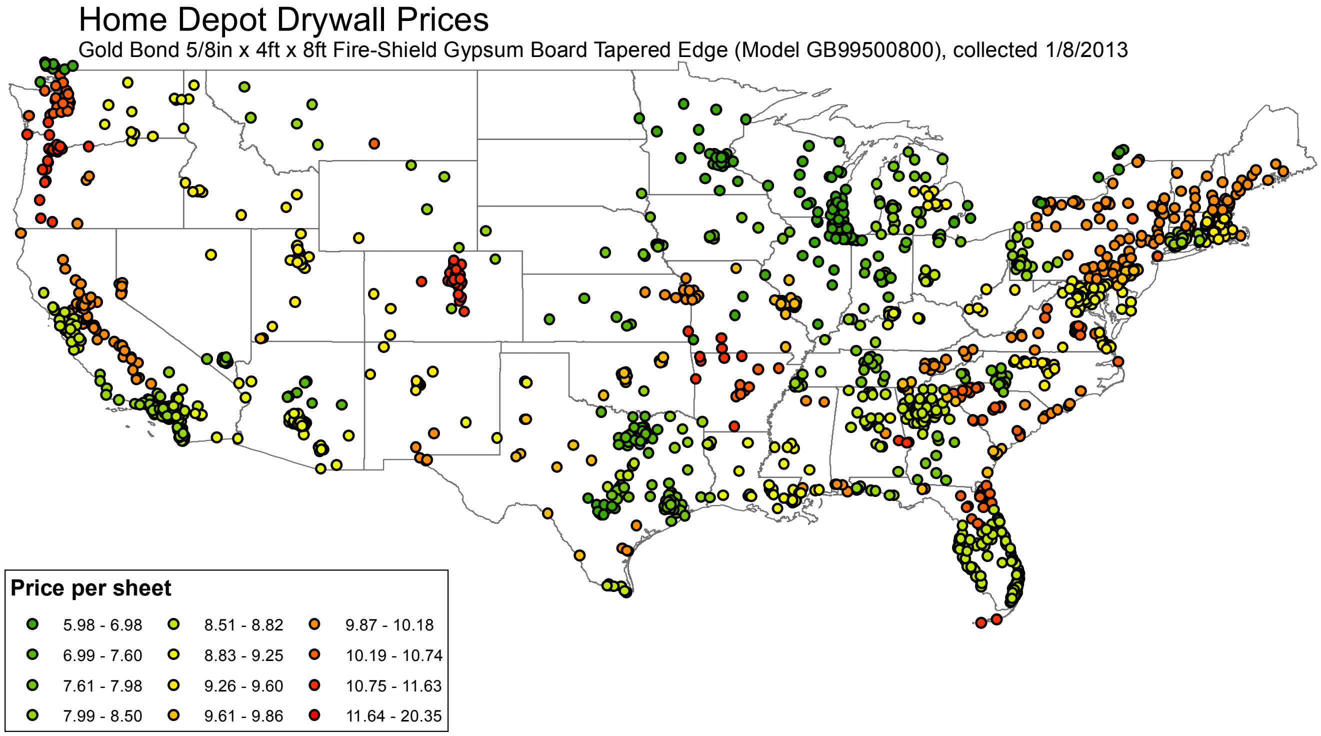


Figure 2.2: Unique prices for regular 5/8” x 4’ x 8’ drywall sheets at Lowe’s

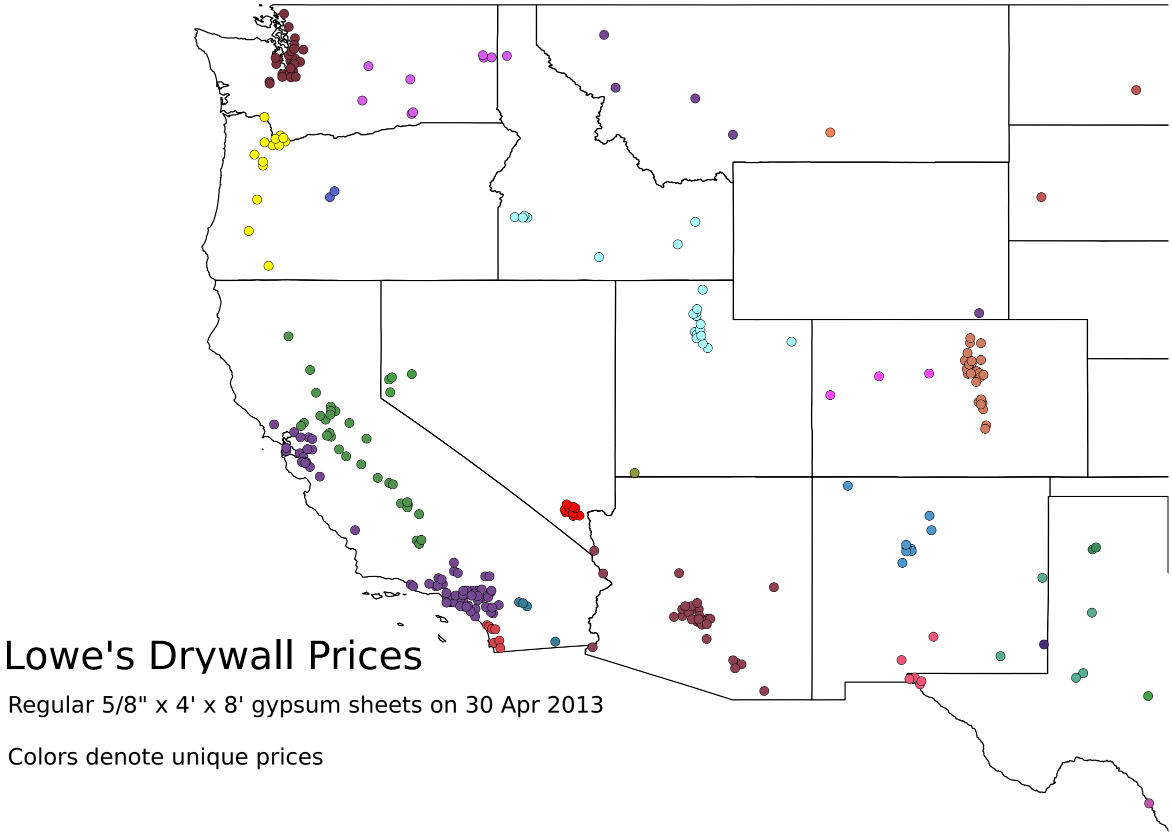


Figure 2.3: Unique prices for regular 5/8” x 4’ x 8’ drywall sheets at Home Depot

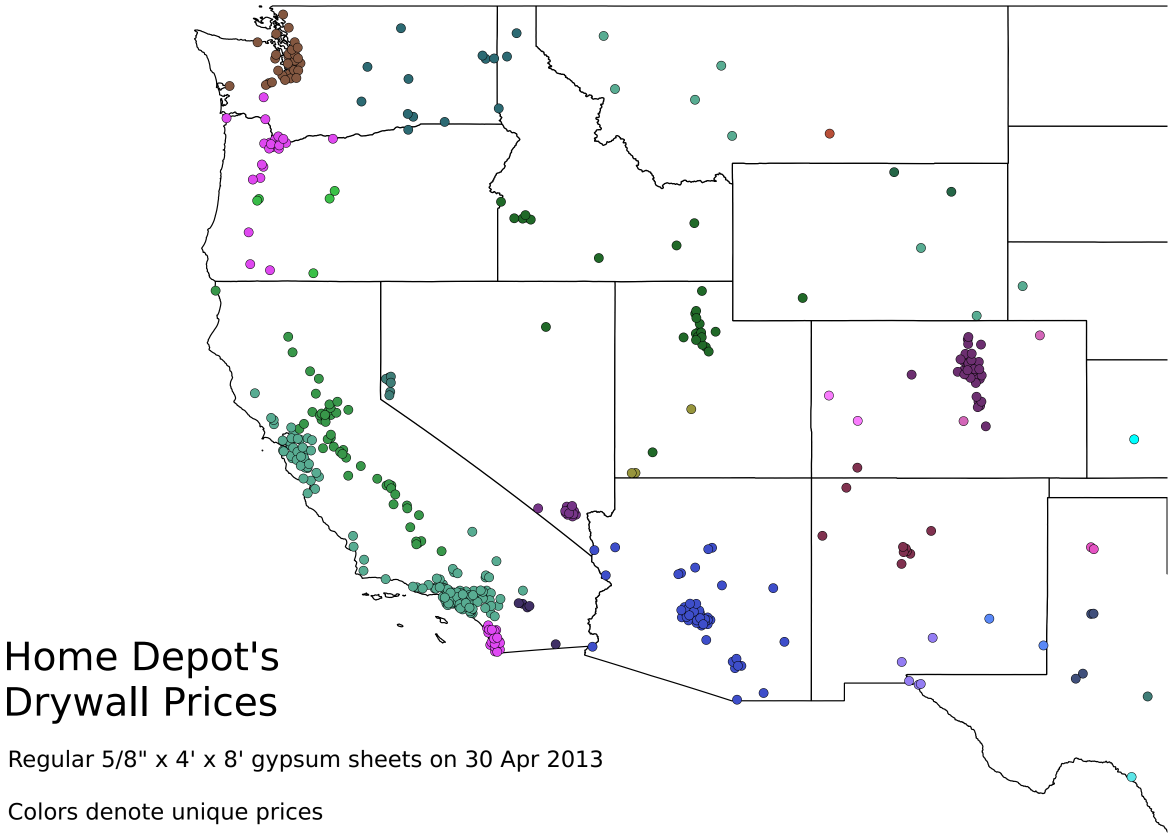


Table 2.1: Size of zones in which all drywall products have common prices

Chain	No. of stores	Pricing zones	Pricing zones with one store	Stores in a zone Mean	Median	Max
Drywall						
Home Depot	1,979	165	35	12.0	6	80
Lowe's	1,714	129	21	13.3	8	94
Hardwood Plywood						
Home Depot	1,979	69	10	28.7	14	232
Lowe's	1,714	247	67	6.9	4	43
Store Brand Interior Paint						
Home Depot	1,979	14	1	141.4	20	743
Lowe's	1,714	1	0	–	–	–

Table 2.2: Prices, cost variables, and competition for three Home Depot stores in one price zone

	Home Depot Stores Logan, UT	Rock Springs, WY	Elko, NV
Prices			
regular 8'x4'x5/8"	$10.98	$10.98	$10.98
mold resistant 8'x4'x1/2"	$11.47	$11.47	$11.47
Distances (miles)			
Nearest Lowe's	1	107	168
Nearest American Gypsum factory	743	821	491
Home Depot Distribution Center	58	177	251

stores. Just as at Home Depot, stores in Utah and southern Idaho post the same price for 5/8" x 4' x 8' gypsum board. However, the prices across chains match as well – $10.98

for 5/8” x 4’ x 8’ gypsum board. Note that Lowe’s pricing zones closely correspond to Home Depot’s pricing zones. In 84.9% of zip codes with both Lowe’s and Home Depot stores, prices for this product match exactly. The success of the law of one price is consistent with theories ranging from Bertrand competition to full collusion, but it is not consistent with Bertrand competition if products are strongly differentiated and costs are different between competing stores. However, in estimating the model, we find that drywall products are highly substitutable within markets.

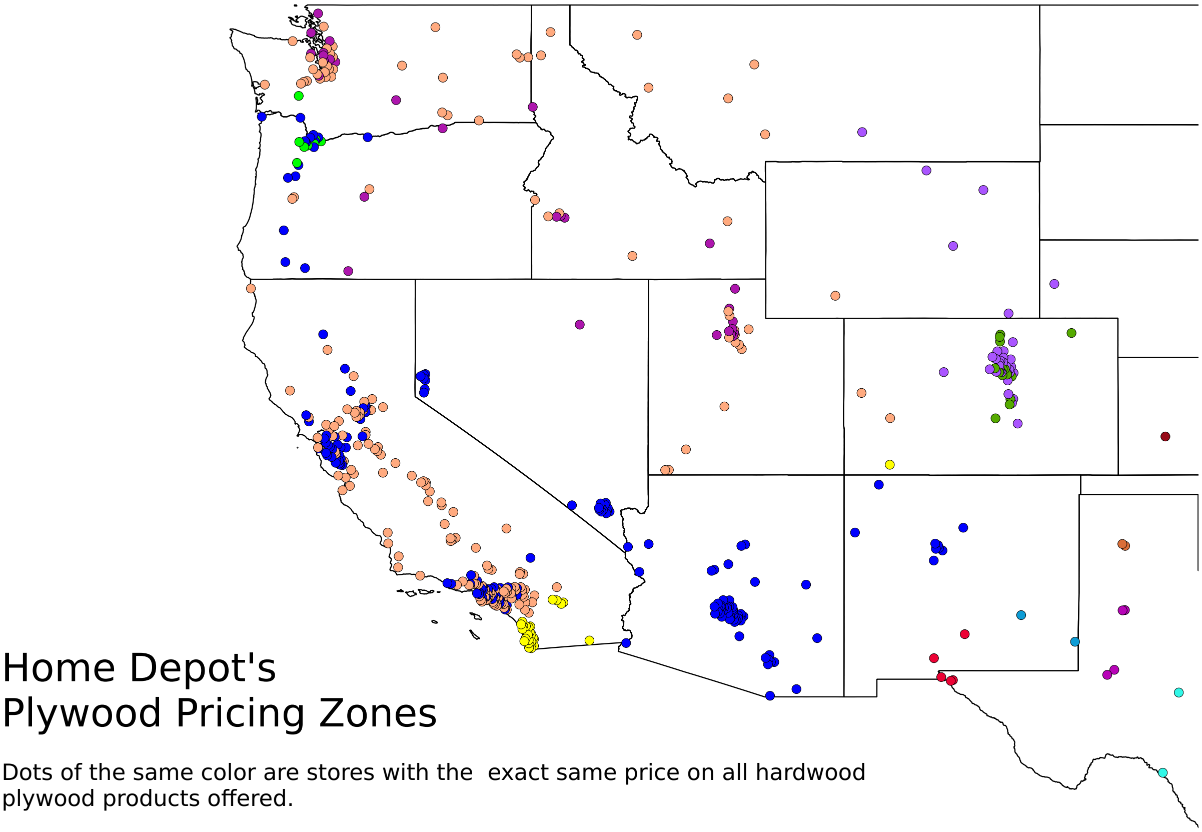
Home Depot and Lowe’s use zone pricing for some other product categories as well; however, the pricing zones differ from observed drywall zones. As documented in Table 2.1, Lowe’s has 247 distinct pricing zones for hardwood plywood whereas Home Depot has 69. At both chains, the boundaries for plywood and drywall pricing zones differ, with dozens of plywood zones intersecting multiple drywall zones and vis versa. Home Depot pricing zones in the Western United States for the paint and lumber products are mapped in Figure 2.5. Note, for example, how Central Oregon is in the same pricing zone as Western Oregon for paint, but not for drywall nor for plywood. Central Oregon shares a drywall zone with Western Nevada, but in plywood it is in one of the two noncontiguous zones that span most of the Pacific Northwest but not Western Nevada.

Interestingly, zones across chains are not the same for all product categories. For example, Home Depot has 14 zones for its in-house paint brand Behr whereas Lowe’s operates uniform pricing for its in-house paint brand Valspar. Having different zone structures by product category is not surprising given that retailers have separate marketing managers for different product categories. The size of zones may depend upon demand elasticities for the product, spatial variability in demand and costs, and the manager-specific costs of interpreting demand data, setting zones, and implementing the pricing program.

Data for Model Estimation

To investigate the spatial menu costs associated with zone pricing in retail drywall, we also create an original data set of prices, sales quantities, and product characteristics for all drywall products at 75 Home Depot stores and 53 Lowe’s stores in the Intermountain West. Each retailer’s web site posts store-specific prices and product availability, along with several product characteristics. We gather a time series of daily inventory levels

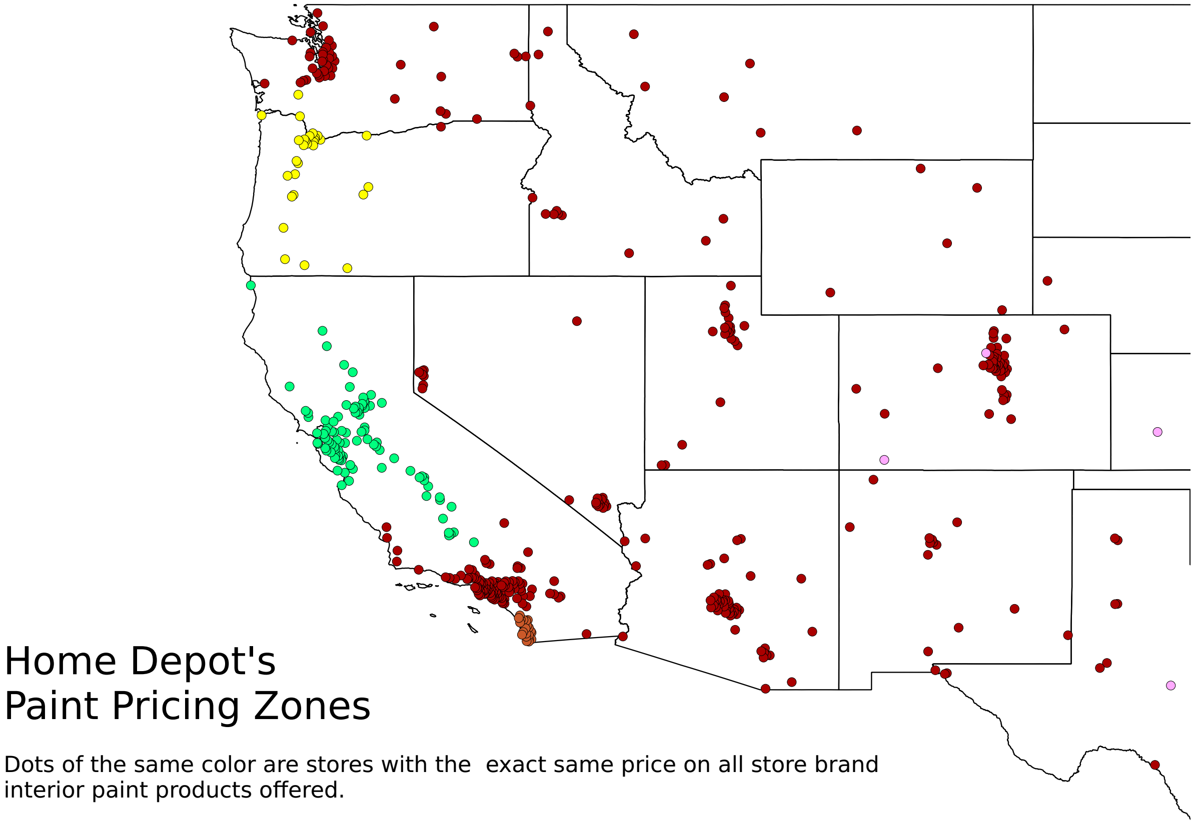
Figure 2.4: Price zones for plywood at Home Depot



Home Depot's
Plywood Pricing Zones

Dots of the same color are stores with the exact same price on all hardwood plywood products offered.

Figure 2.5: Price zones for store brand interior paint at Home Depot



for each product-store combination and use net changes in inventory to calculate sales quantities. Data collection began on 30 January 2013; we now have 113,891 day-product-store observations.

Figure 2.6 maps the stores for which we have obtained quantity data. Our data set includes all stores in Idaho, Montana, New Mexico, Utah, western Colorado, eastern Washington, and stores in adjacent states needed to complete pricing zones. There are locations where only Home Depot operates (for example, Elko, Nevada) and there are locations where only Lowe's operates (for example, Vernal, Utah). Menards, the third largest home improvement warehouse, operates no stores in this region and so can be omitted from the analysis, but it also uses zone pricing for product categories such as drywall and concrete bricks. Of the stores selected, 17 of them are the only store in the 3-digit zip code in which they operate. There are 19 stores located in the 3-digit zip code 840 (neighboring Salt Lake City). Because stores in the Intermountain West are relatively dispersed, distances and labor costs vary, aiding in the identification of marginal costs.

Our sample includes 14 complete Home Depot pricing zones and 11 complete Lowe's pricing zones. The area considered contains several single store pricing zones as well as one of the largest pricing zones. Both the zone boundaries and prices largely match between chains, although Lowe's has a few small zones offering different prices in Cheyenne, El Paso, and Saint George.

We download prices and inventory levels for individual store stock keeping units (skus) and match these to products. For several products, Lowe's lists several brands on their website as different products, but those skus have identical prices and inventory levels that (with a one day lag) coincide perfectly. We eliminate these duplicates. Using manufacturer model numbers, we match products offered by both chains. In all, we identify 31 distinct products. We record the thickness, width, length, mold resistance, and moisture resistance for each product. We do not use brand identifiers because on site visits we found brands frequently mislabeled at both chains.

Net changes in daily inventory levels for each product at each store are used to calculate sales quantities. Decreases in inventory levels give sales quantities. Inventory level increases of more than 20 sheets are classified as deliveries. When deliveries occur, we take the net change in inventory for the day as the volume of the shipment, meaning

Figure 2.6: Stores of interest

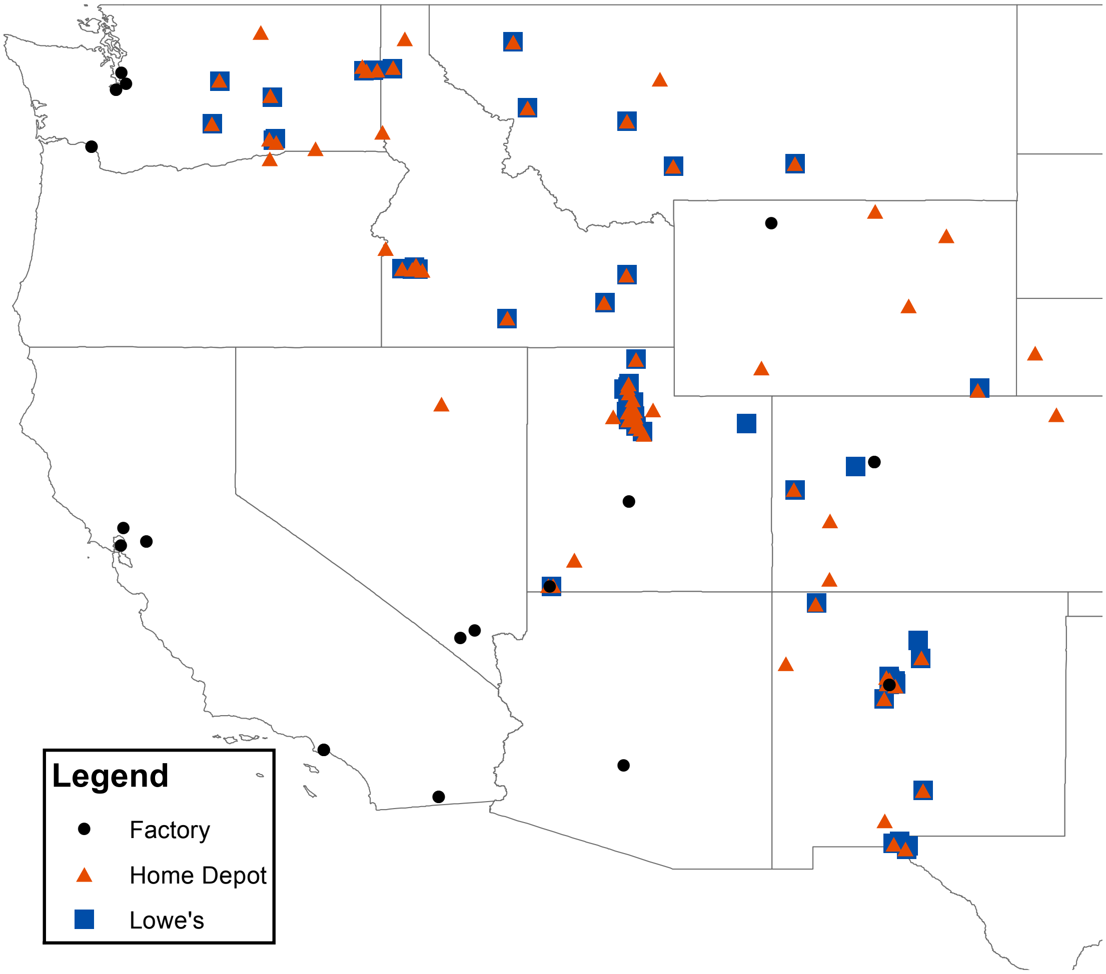


Table 2.3: Summary statistics for the sample

Means	Lowe's	Home Depot
Sales (per product, per day)	6.19 (24.67)	6.97 (19.77)
Delivery size (per product)	200.90 (294.76)	194.11 (216.05)
# Products (per store)	8.60 (1.13)	8.97 (1.15)
Revenue (per store, per day)	$531.87 (381.42)	$651.15 (305.15)
Price (per product)	$10.90 (3.04)	$12.04 (3.78)
obs.	58,100	85,787

we assume no sales take place on delivery days. This systematically under-reports sales, but deliveries occur only every 16 days on average. We do aggregate data which smooths inventory changes. Smaller net increases in inventory levels are counted as returns, or negative sales.

Table 2.3 provides summary statistics for the data sample. On average, daily inventory decreases by 6.2 and 7.0 sheets per product-store for Lowe's and Home Depot, respectively. Because the sales volume is low, we aggregate to the week level in model estimation. While the sales volume appears low, especially given that homes require several hundred sheets of drywall, home builders typically use contract suppliers. We interpret the sales volume to reflect smaller consumer projects, such as wall repair or room remodeling.

On delivery days, typically around a hundred sheets are delivered per product-store. The two chains have similar drywall product selection, offering around eight products per store. The price of drywall within a market ranges from just over \$5 to over \$20 per sheet, depending on the dimensions and features. The total drywall sales revenue for the 128 stores we study sums to \$2.3 million per month.

Both Home Depot and Lowe's operate flatbed distribution centers. These distribution centers provide store locations with lumber and board products.[1] Using distribution center locations, we calculate the closest distribution center to each store, which we utilize in cost estimation. Figure 2.7 displays a histogram of the distances calculated. We find the average distance to stores from distribution centers is 318 miles, with a standard deviations of 190. At the extremes, the closest store to a distribution center is three miles, whereas the greatest distance is just over 690 miles. Distribution centers are usually near large markets and often Lowe's and Home Depot distribution centers are near each other. In our sample region, one big difference is Home Depot's placement of a distribution center in northern Utah, whereas Lowe's nearest distribution center is in southern Nevada. The average weight of a gypsum board exceeds fifty pounds, so we expect distance to be a key driver of costs. Labor costs may also be important. To measure them, we use zip code level wages for home improvement retailers from the Quarterly Census of Employment and Wages.

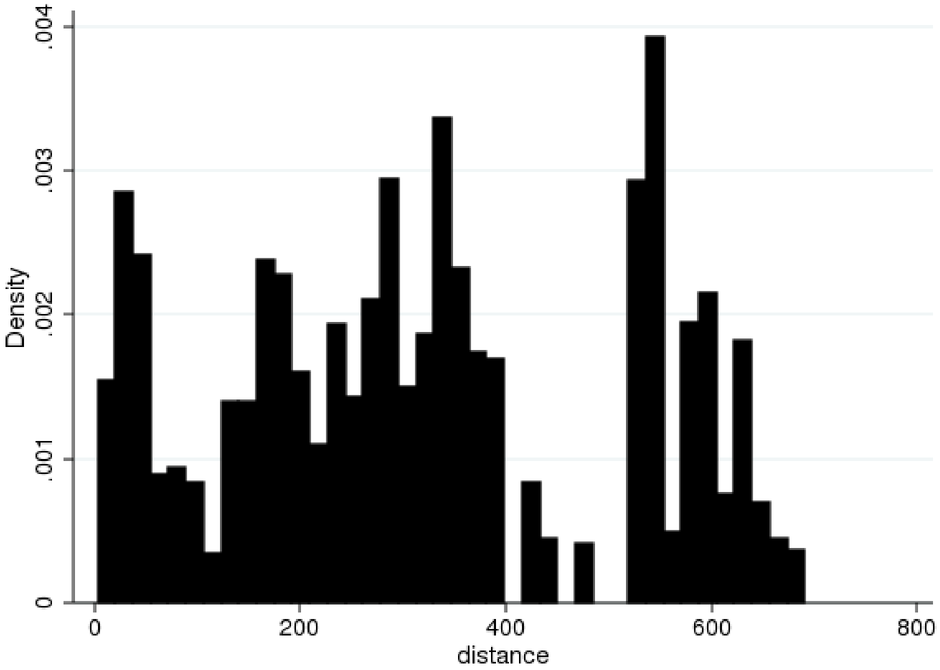
2.3 Model

In this section we introduce the structural model of supply and demand which we use to compare the existing zone pricing regime to alternative pricing policies. The profit gains foregone by not pursuing these pricing strategies will give bounds of the spatial menu costs. To start, we introduce some model notation and define a market. We then describe the discrete choice model used to estimate the demand for consumer drywall. Finally, we specify the pricing problem for firms using zone pricing and show how this constrained problem fits into a more general framework, encompassing both uniform

---

[1] A Lowe's public document states: "FLATBED DISTRIBUTION CENTERS (FDC) - The purpose is to serve Lowe's stores with lumber, plywood, boards, and other building materials that can be forklift loaded onto flatbed trailers."

Figure 2.7: Histogram of distance for each store-product to the closest firm distribution center.



pricing as well as store level pricing.

Notation and Setup

Each of N firms operates a multi-store network where stores are indexed by s. Each store is owned by a firm, superscripted by f and has a location subscripted by ℓ. Firms may operate more than one store at ℓ. Let S_ℓ^f be the set of stores operated by a firm at location ℓ. In total, there are L locations. The concept of zone pricing implies a single price is offered for all products within the zone. Each firm partitions its stores into pricing zones. Let these partitions be denoted by Z^f. For example, if firm f chooses to operate a market level regime, then $Z^f = \{S_1^f, S_2^f, ..., S_L^f\}$. With store level pricing, If $Z^f = \big\{\{s_1\}, ..., \{s_n\}\big\}$, then the interpretation of zones corresponds to stores.

In total, there are J differentiated drywall products, where each store offers a subset of these products each period. Let $J_{s,t}$ be the set of products offered at store s in period t. The product set may change over time due to inventory or the discontinuation or introduction of a product. The discontinuation, introduction, and overall selection of products is not modeled; however, the product set is mostly constant within a zone so there is little evidence of strategic product placement. Given products, product characteristics, and prices, consumers at each location solve for demands. The definition of a market is (ℓ, t).

Demand

Consumers solve a nested-logit discrete choice utility maximization problem. Consider a consumer living at location ℓ. The choice set facing this consumer is the set of products sold by all stores at ℓ. That is, $J_{\ell,t} := \bigcup_{s \in S_\ell} J_{s,t}$. Each period, consumers decide to purchase a single good from the available basket $J_{\ell,t}$ or the outside option. The decision to not purchase a good yields a normalized utility, $u_{i0t} = \epsilon_{i0t}$. Dropping the time subscript, by purchasing a product, a consumer receives indirect utility

$$U_{ij} = x_j\beta - \alpha p_j + \xi_j + \zeta_{ig}(\sigma) + (1-\sigma)\epsilon_{ij},$$

where β are preferences over product characteristics $x_j \in \mathbb{R}^K$, p_{jt} is price, α is the marginal utility to income, and ξ_{jt} is unobserved (to the econometrician) product quality. The composite taste shock, $\zeta_{ig}(\sigma) + (1-\sigma)\epsilon_{ij}$, follows a Type-1 Extreme Value distribution among group g – the nesting variable. The outside good is in its own nest. Note when $\sigma = 0$, the composite error term simplifies to just ϵ_{ij}, which yields the standard logit demand system. As $\sigma \to 1$, there is increasing substitutability of products within nests, and in the limit, when $\sigma = 1$, there is no substitution outside of the nest.

As shown in Berry (1994), given the logit structure of demand, the log difference in market share of good j compared with outside good $j = 0$ equals

$$\ln(\varsigma_j) - \ln(\varsigma_0) = x_j\beta - \alpha p_j + \sigma \ln(\varsigma_j/g) + \xi_j.$$

Here, ς_j/g is the market share of product j within group g.[2] The demand parameters to be estimated are $\theta^D = (\beta, \alpha, \sigma)$. We address the endogeneity of prices in Section 4.1.

Supply

Given the distribution of stores to zones as well as the products offered at stores, firms set prices. Dropping the firm superscript, let z denote an arbitrary zone within Z. Conceptually, if a firm operates four stores across two zones, with a single store in the first zone,

$$\begin{aligned} Z &:= \{z_1, z_2\} \\ &:= \Big\{ \underbrace{\{s_1\}}_{z_1}, \underbrace{\{s_2, s_3, s_4\}}_{z_2} \Big\}. \end{aligned}$$

Zone pricing implies

$$p_{js} = p_{js'}, \forall s, s' \in z.$$

With the example above, the firm uses store level pricing for z_1. However, for all $j \in \bigcap_{i \in 2,3,4} J_{s_i}$, the price is constant for j across z_2. Because the product set is allowed to be different across stores, the definition of a zone implies that if a product is offered at at least two stores within a zone, the price is the same across the stores. The profit of a firm offering product j is

$$\pi_j := \sum_{z \in Z} \sum_{s \in z} (p_{jz} - c_{js}) M_s \varsigma_{js}(X, \xi, \theta), \tag{2.1}$$

---

[2] We use ς_j to denote the purchase probability (market share) instead of the typically seen s_j because we use s to denote a store.

where s has dependence on z. Here, c_{js} is the constant marginal cost associated with offering j at s and M_s is the market size corresponding to the location of store s. Implicitly, only zones and stores that offer j are included in the sum. We assume there are no fixed costs associated with offering products. Firms maximize total profits – the summation of profits over the products offered by the firm.

The profit maximization problem 2.1 is general to include many different pricing regimes. For example, consider a firm utilizing uniform pricing for products. In this case, $Z \equiv z := \{S\}$, so that there is only one zone for the firm and it contains the entire network. In this case, the first sum in 2.1 disappears entirely so that profits are

$$\pi_j := \sum_{s \in Z} (p_{jZ} - c_{js}) M_s \varsigma_{js}(X, \xi, \theta). \qquad \textit{(Uniform Pricing)}$$

With store level pricing, stores and zones coincide so that $s \equiv z$. In 2.1, the second sum disappears and the profits for a firm offering j are

$$\pi_j := \sum_{s \in Z} (p_{js} - c_{js}) M_s \varsigma_{js}(X, \xi, \theta). \qquad \textit{(Store-level Pricing)}$$

Given a set of pricing zones $\mathbf{Z}$, a Bertrand-Nash Equilibrium are prices $\mathbf{p}^* \in \mathbb{R}_+^{|\mathbf{Z}^*|}$, and quantities $\mathbf{q}^* \in \mathbb{R}^{|J_{\mathbf{S}}|}$ such that (1) given pricing zones and competitor prices, each p_{jz}^{*f} solves

$$\pi^{*f} := \max_p \sum_{jz \in J \times Z} \sum_{z \in Z} \sum_{s \in z} (p_{jz} - c_{js}) M_s \varsigma_{js}(X, \xi, \theta; \mathbf{p}_{-jz}) - \mu(\mathbf{Z})$$

where $\mathbf{p}_{-jz}$ denotes all other prices pertaining to the pricing of jz and $\mu(\cdot)$ corresponds to the menu costs associated with a particular zonal structure. For example, with the observed zone structure, equilibrium profits under market level pricing provides a lower bound on the menu costs necessary to move from zone pricing to market level pricing. We assume firms play a game of perfect information in pure strategies.

Unfortunately, the equilibria defined for various $\mathbf{Z}$ are not in general unique. We have found dozens of distinct equilibria for each system of zone partitions we have examined. Caplin and Nalebuff (1991) proves uniqueness for competition within multinomial logit demand systems for single product firms, but its result does not generalize to the multi-product firms we see in our data. In particular, there are equilibria in which firms assign high prices to some products to take advantage of the tail consumers with particularly high ϵ_{ij} draws while shepherding the rest of the consumers into moderately priced products. Different equilibria assign the role of the moderately priced mainstream alternative to different products and vary as to which store engage in this version of price discrimination.

In order to evaluate the menu costs associated with alternative pricing policies we must ensure that differences we find are due to the policies themselves, and not due to switching between vastly different equilibria. We therefore investigate small deviations from the observed equilibrium, allowing firms to switch to market level pricing one product at a time. By only allowing for adjustments in prices of a specific good in a single market, the result of Caplin and Nalebuff (1991) does guarantee unique equilibria. Hence, in our definition of a Bertrand-Nash Equilibrium, we condition on the zone structure instead of having firms choose zones. These experiments are reported in Section 5. In the Appendix, we explore a move to market level pricing for all products simultaneously. We discuss a metagame of zone choice and pricing, and calculate menu costs after implementing a selection mechanism on equilibria.

2.4 Estimation and Results

We proceed by estimating the demand parameters which enter shares ς as well as marginal costs. We estimate the model in two stages. First, we estimate the demand parameters of the nested logit model. Given estimates of the demand parameters, we solve for marginal costs assuming firms are competing in a Bertrand game of zone pricing. As zone pricing is a consequence of solving a constrained optimization problem, we cannot invert the first-order conditions to back out marginal costs. Instead, we parametrize the cost function, and simultaneously recover marginal costs and cost parameters using mathematical programming with equilibrium constraints (MPEC). With our estimates, we calculate observed profits of the current zone structure. Given the presence of multiple equilibria, we utilize observed prices to calculate current zone profits.[3]

---

[3] By solving for equilibrium zone prices given observed prices as starting values, we obtain equilibria prices quite close to observed prices. The geometric fit between the two is 93%. This is discussed further in the Appendix.

Demand

We invert market shares, as shown in Berry (1994), to obtain $\ln(\varsigma_j) - \ln(\varsigma_0) = \delta_j$, where δ_j is the mean utility from purchasing product j. To account for the endogeneity of unobserved product quality being correlated with price, we pursue a fixed effects approach. We separate unobserved product quality as

$$\xi_{jt} := \xi_j + \xi_t + \Delta\xi_{jt},$$

where we assume $\Delta\xi_{jt}$ is uncorrelated with price and observed product characteristics. We estimate ξ_t as a market-time fixed effect. As a robustness check, we compare this method with an instrument variables approach, where we instrument price by using a Hausman instrument – average prices for a given product in other markets where the product is offered. For product characteristics, we use the dimensions of the gypsum board (length, width, height), and whether the drywall is mold resistant and/or moisture resistant.

We use a product-store hierarchy for nests within markets. The top level of nests denotes the various product types available in the market (ℓ, t). We define product types as the grouping of product dimensions. The second nest comprises the various stores in the market that sell that particular product type. The interpretation of σ in this nesting structure is the degree of substitutability of a product type across the various stores at location ℓ at time t. We specify the nesting structure this way because of conversations we have had with home builders, who say that by far, the most important characteristic of drywall is the size, particularly the thickness. Specifically, almost all walls use 1/2-inch boards, but 5/8-inch is necessary on fire walls, such as the walls separating the interior from a garage. If size is the most important characteristic, we would expect that consumers substitute to other stores in the market for a particular drywall type instead of substituting to a different size sheet at the same store. Hence, our nests are at the product type level instead of the store level. We expect to (and do) estimate σ close to one, which suggests that a particular product type is highly substitutable across stores within a market.

Identification for parameters in this stage results from the observed purchases of consumers given their choice set, following the standard revealed preference identification for discrete choice demand systems. In every market, all products and their associated prices and characteristics, are known. Product indicators and characteristics are constant across all markets. The other products offered vary by market and relative prices vary by zone. The response of sales quantities to these different relative prices and product offerings identifies the price coefficient α. Preferences for observed and unobserved product characteristics are revealed through market shares.

We estimate several demand specifications. First, we set $\sigma = 0$ in the nested logit model so that the nests do not matter. This results in the classic logit demand system. We estimate this specification assuming prices are exogenous (OLS) and then use a fixed-effects approach to address the endogeneity of prices. We then estimate σ along with the demand parameters, again assuming prices are exogenous, and then accounting for endogeneity. Given the nesting structure, all observed products characteristics within a nest are identical, except price (across the stores in the market). Hence, for the nested logit model using instrumental variables, we use store fixed effects to address the endogeneity problem on the group shares. For the nested logit model using fixed effects, the endogeneity problem on group shares is already addressed by having the fixed effects be product-store ("j") specific.

We aggregate daily sales so that t denotes a two week period because many drywall products have few daily sales. Observations with zero sales must be dropped, because $\ln(\varsigma_{jt})$ would be undefined[4] At the biweekly level, 5.70% of products exhibit zero sales. Aggregating data across time does not introduce as much measurement error as it might in other applications. Our product characteristics are all time-invariant. Measurement error may be induced because of changes in product prices. However, in the data, prices change only rarely. Over 127 days of data collection, only 9.4% of product-store combinations exhibit price changes. Of the products to see price adjustments, 88% (77 products) of them experience a single change and 12% (11 products) see two price adjustments.

In order to complete the demand system, we need to specify market size. We define a market to be a Core Based Statistical Area (CBSA), two week period. For stores not located within CBSAs, we set the market to be the county in which the store resides. With this interpretation of markets, each location ℓ typically has several stores from

---

[4] See Gandhi, Lu, and Shi (2013) for estimating discrete choice demand systems with products that exhibit zero sales.

both Home Depot and Lowe's. Further, given the structure of pricing regions by both firms, regions overlap into several markets. In El Paso, TX, both Home Depot and Lowe's have multiple zones within the same CBSA – two zones divide the city. We take market size to be proportional to the 2010 CSBA population.[5]

The results of the demand estimation using a fixed effects approach appear in Table 2.4. For comparison, and as a robustness check, Table 2.5 reports demand estimates using an instrumental variables approach. Across all specifications, all coefficients have expected signs and are significant across specifications. We estimate that consumers prefer larger drywall sheets and mold resistant panels. The unreported coefficients on drywall thickness are reasonable and show that industry standard 1/2-inch panels are much more desirable than all other thicknesses. We estimate that consumers are price sensitive, with the marginal utility of income -0.375 in the fixed effects, nested-logit model (Model 4). Our estimates on price sensitivity do not become more negative when accounting for the potential endogeneity between prices and unobserved drywall quality, which is not what we would expect as high priced items are typically assumed to be positively correlated with unobserved quality. This is true for both the fixed effects and instrumental variables approach. Our interpretation on the price coefficient across specifications is that given any two drywall products with identical (observed) characteristics but different prices, consumers would gravitate towards the cheaper good as the other (unobserved) characteristics are not worth the additional expense. Hence, in our setting, the correlation between unobserved quality and price is negative leading to price sensitivities closer to zero after accounting for endogeneity. We estimate the mean own price elasticities to be -16 and -18 for the fixed effects and instrumental variables nested logit models, respectively. These values are large in magnitude but not unreasonable given the high substitutability of drywall products, especially within nests. For the final specification (4), we obtain industry elasticities of -0.03 to -0.04 depending on the market. Finally, we estimate the coefficient reflecting substitutability within nests to be high, at .830 in the last specification. This suggests that consumers would rather drive to another store within the market to buy a particular drywall panel

---

[5] For observations not within CBSAs, we take the population to be proportional to the 2010 Census county population.

than substitute to a different size.[6]

In the following analysis, we use Model (4) – the nested logit model with fixed effects – as our model of consumer demand. Our results are not sensitive to this choice as the nested logit model with instrumental variables yields quantitatively similar answers. Aggregating the data to just the week level also yields similar results.

Recovering Marginal Costs

Marginal costs can typically be recovered using the demand estimates and the first order conditions of each profit maximizing firm. In the single good, market level pricing case, Lerner's index is inversely proportional to the own-price elasticity. With observed prices and an estimated demand elasticity, the marginal cost is identified. A multi-product analog, as seen in Nevo (2001) and Petrin (2002), can be used when firms set prices for all products at the market level. However, our environment differs from the standard setting as prices are not set at the market level. Rather, they are set uniformly, possibly across several markets.

When firms are constrained to set a single price for a product in a whole zone, only one first order condition per zone is present for each product. Given the firm's profit equation in 2.1 under zone pricing, each price p_{jz} is obtained from solving

$$\frac{\partial \pi}{\partial p_{jz}} = \sum_{s \in z} \sum_{i \in J_s} (p_{iz} - c_{is}) M_s \frac{\partial \varsigma_{is}}{\partial p_{jz}} + \sum_{s \in z} M_s \varsigma_{js} = 0. \tag{2.2}$$

The first-order condition for each price contains marginal costs for all stores within its pricing zone. The supply system yields $|J \times Z|$ first-order conditions of the form in 2.2. However, there are $|J \times S|$ marginal costs to identify and $|S| > |Z|$. As there are more marginal costs than first-order conditions, no set of first-order conditions can be directly solved to recover marginal costs.

To make progress in recovering marginal costs, we first parameterize costs as

$$c_{jst} = a_j + w_j' \kappa + \nu_{jzt},$$

where a_j is a fixed effect for product j and w_j contains both the distance (d_{js}) traveled for product j to store s as well as the local wage. We use home improvement retail

---

[6] Our model of demand assumes it is costless for consumers to travel to stores within a market.

Table 2.4: Demand estimation results

	(1) Logit	(2) Logit FE	(3) Nested logit	(4) Nested Logit FE
price	-0.612***	-0.258***	-0.663***	-0.375***
	(0.0197)	(0.0241)	(0.0132)	(0.0169)
area	0.134***	0.0693***	0.160***	0.0978***
	(0.00486)	(0.0124)	(0.00327)	(0.00853)
moldresist	1.229***	1.538***	1.420***	2.038***
	(0.0801)	(0.448)	(0.0537)	(0.321)
chain	0.970***	0.431***	0.330***	0.193
	(0.0408)	(0.0365)	(0.0282)	(0.139)
σ			0.921***	0.830***
			(0.00995)	(0.0107)
constant	-6.172***	-8.245***	-4.402***	-8.389***
	(0.109)	(0.438)	(0.0752)	(0.395)
Thickness indicators	Yes	Yes	Yes	Yes
e^D_{jj}	-6.396	-2.790	-59.056	-16.510
	[-10.045,-2.691]	[-4.415, -1.158]	[-141.914, -2.951]	[-35.482, -1.683]
R^2	0.250	0.592	0.663	0.810
N	6985	6985	6985	6985

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.5: Robustness check: Instrumental variables demand results

	(5) IV Logit	(6) IV Nested Logit
price	-0.806***	-0.880***
	(0.0273)	(0.0162)
area	0.177***	0.204***
	(0.00647)	(0.00384)
moldresist	1.748***	1.968***
	(0.0952)	(0.0566)
chain	0.962***	1.760***
	(0.0411)	(0.218)
σ		0.595***
		(0.0145)
intercept	-5.601***	-5.910***
	(0.133)	(0.319)
Thickness indicators	Yes	Yes
e^D_{jj}	-8.395	-18.387
	[-13.126,-2.844]	[-35.684,-3.006]
N	6985	6985

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

wages at the zip code level (NAICS 44411). Finally, ν_{jzt} is an unobserved (to the econometrician) cost shock at the product, zone, time level. Let $\theta^S = (\mathbf{a}_j, \kappa)$. Because the cost shock is at the product level, we cannot manipulate 2.2 to recover costs directly; however, given an objective function on ν, we can simultaneously recover marginal costs and the parameters governing costs. Instead of using a nested fixed-point approach, we

Table 2.6: Cost Estimates

	Point Estimates	Std Error
dist (κ)	0.00052	(0.00016)***
	Product Fixed Effects	
a_1 : 8.384	a_{11} : 11.631	a_{21} : 12.609
a_2 : 7.561	a_{12} : 12.154	a_{22} : 11.323
a_3 : 2.088	a_{13} : 10.670	a_{23} : 14.052
a_4 : 7.137	a_{14} : 10.900	a_{24} : 18.613
a_5 : 7.947	a_{15} : 12.395	a_{25} : 13.860
a_6 : 7.948	a_{16} : 12.063	a_{26} : 9.760
a_7 : 9.307	a_{17} : 8.921	a_{27} : 9.989
a_8 : 14.592	a_{18} : 11.042	a_{28} : 15.298
a_9 : 7.137	a_{19} : 11.180	a_{29} : 10.236
a_{10} : 12.990	a_{20} : 13.269	a_{30} : 14.363
		a_{31} : 14.873

Zone clustered standard errors. All FEs significant at 1%

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

proceed with using mathematical programming with equilibrium constraints (MPEC) as seen in Su and Judd (2012). Forming moment conditions on ν directly, along with the optimality conditions from the firms' zone pricing problems, completes the mathematical program.

Firms want high prices at stores with high costs and stores with monopoly power. An unconstrained optimum price for a competitive, low-cost store would be lower. When a firm has both types of stores in the same price zone, the optimal price balances these considerations using 2.2. Only the zone price, market power (through estimated price elasticities), market size, and a few cost variables are observed, but the zone price reveals information on the marginal costs for its stores. Identification effectively comes from how weighted averages of store cost variables are correlated with the zone price. For example, if zones full of competitive stores far removed from a distribution center has a high zone price, then distribution center distance is an important driver of costs.

Our simplification in making the unobserved error term ν be product, zone, time specific instead of product, store, time specific results in the dimension of the error term being equal to the number of equilibrium conditions. Without this assumption on the cost shock, we would have an unidentified system. Although restrictive, we do account for transportation costs and wages at the store level.

The objective comes from moment conditions on ν. Let $X := [\mathbf{a}_j, W]$ be the matrix of covariates on costs. The method of moments estimator is derived from $\mathbb{E}[X'\nu] = \mathbf{0}$, leading to the sample analogue

$$g_j(X, \theta^S) = \frac{1}{N}\sum_{i=1}^{N} x_{ji}\nu_i = 0.$$

Letting $\text{FOC}(\theta^S, \nu)$ denote the set of equilibrium conditions characterized by the first-order conditions of the firms' problems, the MPEC program is to solve

$$\begin{aligned} &\min_{\theta^S,\nu} g(X, \theta^S)' g(X, \theta^S) \\ &\text{s.t. } \text{FOC}(\theta^S, \nu) = 0. \end{aligned}$$

In estimating costs, we obtain a negative, but insignificant coefficient on wage. We drop wage from the model and proceed with estimating product fixed effects and the coefficient on distance. The remaining coefficients are very similar to the model with wages included. Cost estimations are reported in Table 2.6. We estimate the parameter on distance (per mile) to be \$0.00052. On average, transportation from the distribution center contributes \$0.20 to the cost of a drywall sheet. We find transportation costs for different stores range from \$0.002 to \$0.46. The coefficient on distance is lower than other estimates of transportation costs in similar settings. Miller and Osborune (2011) estimate a transportation cost \$0.30/ton mile for Portland cement. The equivalent cost for a fifty pound sheet of drywall would be \$0.0075/mile, which is over eleven times larger than we find. If, however, other products shipped to stores on flatbed trucks need frequent deliveries, perhaps much of the drywall inventory is shipped on trucks with spare capacity. Indeed, we find that the deliveries for drywall are less than the full capacity of a flatbed trailer.

Observed Pricing Regime

With observed prices and the marginal costs we estimate, we calculate the sales weighted average markup on a sheet of drywall to be \$1.11. With an average price of a sheet of drywall at \$10.22, we find the margin on drywall to be around 11.0%. We estimate profits on drywall for the stores of interest to be about \$29 million annually. Table 2.7 details equilibrium zone pricing profits by chain and competition type. Only 3 of the 53 Lowe's stores in our region are in markets where Home Depot is absent. Home Depot on the other hand has 20 or their 75 stores in markets without competition from Lowe's. These twenty account for 38% of Home Depot's revenue. Interestingly, 15 of the 20 Home Depot monopoly stores are in pricing zones with stores that do face competition from Lowe's. A higher price that would extract the most profit in the monopoly markets must be balanced by a lower price needed to maintain market share in competitive markets. Because several of the monopoly store are in large, lower cost zones, average prices for monopoly markets are slightly below average prices overall. As a consequence, the current zone structure greatly limits Home Depot's effective market power. Indeed, we estimate Home Depot only obtains 16.48% of its profits from monopoly stores.

Figure 2.8 plots a histogram of observed profits, by chain, aggregated by store. The histograms show there is considerable variation in profits across stores. In this region, Home Depot distribution centers are closer to more stores, so our estimates generally find Home Depot stores to have lower costs and higher profits than Lowe's stores. We estimate a majority of the Lowe's stores have less \$1 million in annual profits for drywall, with the maximum profits being nearly \$2 million annually. On the other hand, Home Depot operates a few stores that exceed \$2 million in annual profits. Both chains operate stores with nearly zero profits from drywall sales.

2.5 Alternative Pricing Regimes

Here we calculate optimal pricing policies by changing the zone structure of firms. Given the presence of multiple equilibria, we investigate moving one product from its current zone pricing scheme to market level pricing. In this exercise we hold all other product prices constant at their observed levels. This guarantees a unique equilibrium for the exercise. We first consider holding competitor prices fixed. In a second exercise, we

Figure 2.8: Histogram of observed profits, by chain, and aggregated to the store level.

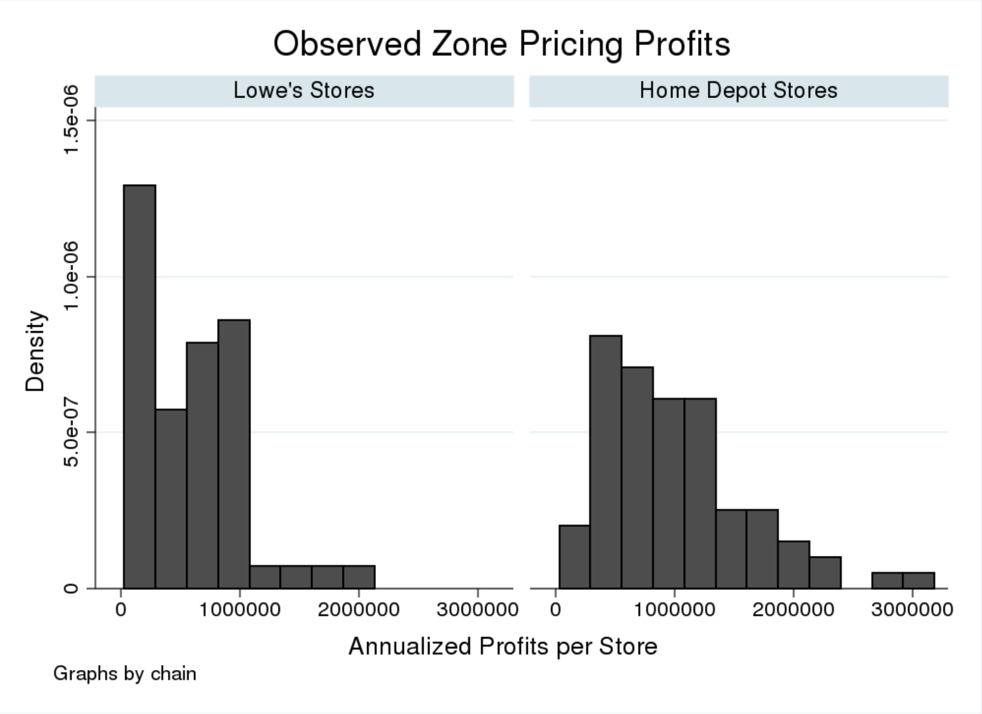


Table 2.7: Current Profits by Chain and Market Type

	Lowe's		Home Depot	
	% of π	Annual π	% of π	Annual π
Monopoly	3.28%	\$273,462	16.48%	\$3,356,029
Duopoly	96.72%	\$8,057,492	83.52%	\$17,007,531
Total Annual π		\$8,330,954		\$20,363,560

Duopoly means there is a competitor store in the market (CSBA). Profits are annualized.

allow both chains to switch to market level pricing for a single product. Previous work on zone pricing has not taken into account this competitor response, and we highlight that this leads to an overstatement in profit gains by 33% in retail drywall. We discuss how profit changes in either experiment relate to various interpretations of spatial menu costs.

Unilateral Single Product Deviations

We first let one chain switch to market level pricing for one product while holding fixed prices on all of its other products and prices on all of its competitor's products. Market by market we calculate new prices that maximize total firm profits. For this exercise, profit gains must be nonnegative as the zone price is in the choice set and all other prices are unchanged. In some sense, the profit gains here represent a lower bound on menu costs. If post-menu cost profits were higher with market level pricing given competitor prices, then the observed zone pricing would not be a best response to those competitor prices and hence not an equilibrium strategy. This also parallels the approach of the prior literature, which estimate profit gains from store level pricing implicitly keeping competitor prices fixed by not modeling competition.

A zone price must balance the profit gains available by raising the price at stores facing high residual demand or high marginal cost with the losses that would follow from overcharging at stores in the zone where marginal costs are lowest or residual demand is most elastic. The optimal market-specific prices are spread both above and below the zone price. In our sample, when one firm can adjust one product-market price, we find prices in 62% of product-store observations increase and decrease for 38% of product-store observations. Together the sales weighted average price increases by \$0.12. As expected, prices in monopoly markets increase. Figure 2.9 provides a density plot of price deviations by market competition for Home Depot stores. As expected, the average price in monopoly markets increases \$1.63. For some products in monopoly markets, prices increase by over \$3 or by nearly 25%. In contested duopoly markets, the average price decreases by \$0.08. While the average deviation in duopoly markets is negative, Figure 2.9 shows that for some products Home Depot increases prices in duopoly markets by over \$2. Some of these price increases occur because Lowe's usually offers the same product but does not stock it in that particular market, giving the Home

Depot store some market power for that product.

For monopoly stores in their own pricing zone, the zone pricing problem and this exercise coincide. However, since costs are not perfectly measured, observed zone prices are not necessarily equal to optimal zone prices. Hence, when solving for optimal prices for products in monopoly markets and own zones, we find small price adjustments to solve the first-order conditions on the firm's problem. Therefore, we report nonzero profit gains for these products. These errors are lumped together with the menu cost. An alternative exercise would begin with a simulated equilibrium instead of observed prices. There are multiple pricing equilibria for the current zone structure, and while some are very close to observed prices and quantities, we would need to dictate an equilibrium selection rule. Here we begin with the equilibrium that has been observed and proceed with small changes that would force the selection of a new equilibrium.

When summing the profit gains from all the single product deviations, we find market level pricing to give profit gains of around 3.3% of revenue, or 33% of current profits. This equates to nearly $6.7 million in additional industry profits annually for the 128 stores of study. We calculate the average annual gain to be $10,430 per store for Home Depot and $11,617 per store for Lowe's.

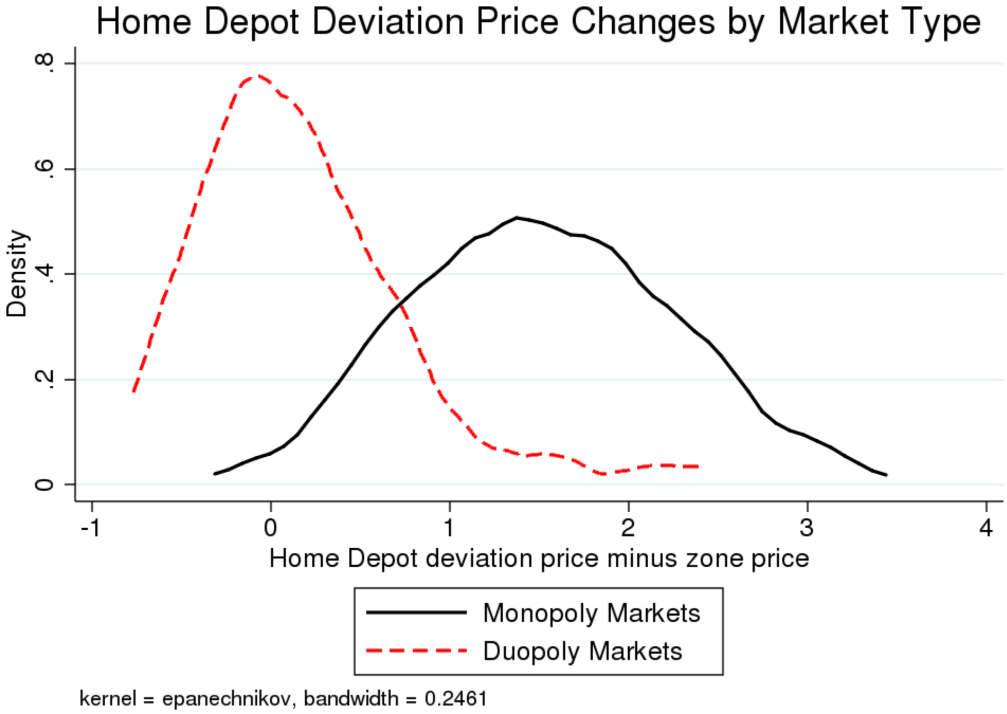
Our finding are consistent with the previous empirical studies of zone pricing and with elementary price discrimination theory. Without competitive effects, offering separate prices in each market will increase profits while lowering prices for some consumers and raising them for others.

Single Product Market Equilibria

We next allow competitors to respond with market level pricing of their own, again in one product. For that product, we find an equilibria in which both chains set a market-specific price that are best responses to the others price. The equilibrium prices we find have the same general features as in the first exercise. In duopoly markets prices decrease $0.09 from the observed zone price. The price increase in monopoly markets are the same, since there the equilibrium price and the unilateral deviation price are identical. In 55% of product-store observations prices increase.

Prices that were lowered in the unilateral deviation are decreased further market prices are decreased further in equilibrium in 74.4% of observations in duopoly markets.

Figure 2.9: Density of Home Depot price changes when a chain adopts market level pricing holding competitor prices fixed. The plot shows that in monopoly markets, prices only increase whereas in duopoly markets, prices both increase and decrease.



In the Salt Lake City market, for example, 4'x8'x5/8" regular drywall is sold for \$10.98 when firms use zone prices. Both Lowe's and Home Depot have monopoly stores and distant, higher cost stores in the same price zone, and their zone prices reflects this. If all other prices were fixed, Home Depot would set a market level price of \$10.70. Lowe's would charge \$10.74. In equilibrium, Home Depot has to compete not against a fixed \$10.98 Lowe's prices, but against a more competitive Lowe's price. Likewise, Lowe's must respond to a lower Home Depot price. Home Depot's equilibrium price is reduced to \$10.63. Figure **??** compares the prices (relative to observed levels) for Home Depot when Lowe's prices are held fixed and when Lowe's can respond with its own market-specific price.

In a few markets, one firm's zone price falls below their optimal price, while the other firm's is above their optimal price. This can come from cost differences stemming from the dissimilar placement of distribution centers or from one firm's zone containing more monopoly stores that push up the optimal zone price. In these markets, equilibrium considerations moderate the price movements seen in the unilateral deviation experiment. In some high cost markets, a firm in the unilateral deviation experiment will double its price, forfeiting the mainstream market to capture consumers with strong product-specific preferences. If the rival chain raises its price, its equilibrium price will be lower than the unilateral deviation, as returning to something closer to normal competition becomes more appealing. Under certain conditions, to ignore competition would cause overstatement of the effect of moving to market level pricing, but in our region understatement was more common.

Because competitive effects are absent in monopoly markets, the effects that competition magnifies are mostly price decreases in duopoly markets. Because industry demand is so inelastic, these price decreases reduce profits. In total, equilibrium profit gains are 32.9% lower than in the move to market level pricing holding competitor prices fixed. The sum over all products of moving from a zone pricing equilibrium to a market pricing equilibrium are 2.2% of current revenue. The profit gains are almost evenly split between Home Depot and Lowe's in total; per store Lowe's gains more: \$8,879 to Home Depot's \$6,665.

In the price discrimination theory literature, the profit effects of increased ability to price discriminate depend on the industry and cross-price elasticities of demand. Drywall industry demand is nearly inelastic and many products are close substitutes, bringing it close to the conditions in Holmes (1989) where discrimination yields lower industry profits than uniform pricing. Yet, in our counterfactual experiment, industry profits increase. Some of this is driven by the presence of monopoly markets, but even in duopoly markets profits increase on net. Furthermore in every zone, industry profits increase under market level pricing, although Lowe's profits decrease in a few zones.

If firms can credibly commit to their pricing zones, the (pre-menu cost) profit gains calculated here are lower bounds on the relevant menu costs. If the menu costs for market level pricing are more than a chain's profit gains, it would prefer to stay in the zone pricing equilibrium rather than invest in market level pricing. The magnitude of bounds we calculate here could be consistent with a managerial source to the spatial menu costs. Retail segments with more elastic industry demand or with less substitubility between competitors should have the higher profit gains. If they are enough to spatial menu costs, retailers market level pricing or smaller pricing zones.

2.6 Conclusion

In this chapter, we document the prevalence of zone pricing in home improvement retail stores. Although product categories such as drywall and lumber have sizable price variation nationally, within regions there can be no price variation within firm. The size of zones, including the number of stores and the number of markets per zone, varies from product category to product category. We find some product categories with hundreds of zones, and for other categories, a single firm, or uniform pricing, is pursued by a firm. Having different zone structures by product category is not surprising given that retailers have separate marketing managers for different product categories. The choice of the zone structure reveals how firms balance discrimination and competition across markets. We postulate that the use of zone pricing, instead of a finer grade of pricing, such as by market or by store, is the result of firms facing a friction – "spatial menu costs". These spatial menu costs have induced firms to set a constant price over multiple markets.

To provide a measure of the spatial menu costs needed to rationalize the use of zone pricing, we estimate a structural model of consumer demand on a detailed data set of retail drywall. We find that consumers consider the products of competing chains to be close substitutes, but the industry elasticity for drywall is inelasitic. Assuming firms are engaged in Bertrand price competition, we back out marginal costs to find that transportation costs are a small, but significant component of costs.

Given our estimates on supply and demand, our menu costs are calculated by comparing the observed profits in zone level competition to the equilibrium profits in which firms adopt market level pricing. Since firms are offering multiple products, priced uniformly across several markets, multiple equilibria exist. To obtain menu cost estimates, we investigate small deviations in the current zone structure which result in unique counterfactual equilibria. Finding the equilibria for these alternative pricing regimes yields a lower bound on the spatial menu costs at 2.2% of current revenues or 22% of observed profits. While the previous literature on zone pricing has determined menu costs may be large, these articles have not taken into account the competitive interaction of firms. We find that ignoring competitive effects by fixing opponents prices implies much larger gains from market level pricing which overstates the spatial menu costs by upwards of 32%, at 3.3% of observed revenues.

Spatial menu costs force firms into a price zone system that prevents them from abusing their market power. In an industry like drywall with high transportation costs and inelastic demand, menu costs and zone pricing protect consumers in monopoly markets. The elimination of the menu costs would prompt a new level of strategic competition in the duopoly markets, but according to our estimates would still leave the retailer more profitable.

Essay Appendix: Metagame Analysis

In this section we explore the metagame in which firms choose to adopt zone pricing or market level pricing. There are millions of zone combinations so we only explore the option of selecting the current zone structure for each firm, or a move to market level pricing for the entire network. We calculate the Bertrand-Nash Equilibrium (BNE) for four pricing regimes: Lowe's and Home Depot keep their current zone structure, Lowe's moves to market level pricing and Home Depot keeps its current zone regime, Home Depot moves to market level pricing and Lowe's keeps its current zone structure, and finally, both firms move to market level pricing. Prices adjust for all products in all periods.

As previously noted, there are multiple equilibria for each of these pricing regimes. We utilize a selection mechanism on the number of products with low sales. Due to the logit error term, high prices yield marginal sales, but the firm may choose to set very high prices so that consumers substitute to other products with better margins. Of the equilibria found, we select the equilibrium for each scenario that has the lowest number of products priced sufficiently high as to yield marginal sales. For example, in solving for equilibria based on the current zone structure, there are equilibria in which the price of a product is such that sales are close to 10^{-6} sheets per week. We sum up the number of product-store combinations in which this occurs, and select the equilibrium with the lowest number. We gauge the performance of this selection mechanism by comparing the observed zone equilibrium with the calculated zone equilibria. The lowest number of observations with marginal sales is 306. The median difference between equilibrium and observed zone prices is \$0.007, and the mean difference is \$0.11. Lowe's observed annual profit is \$8,330,954 for the 53 stores in the sample. With our selection mechanism, we calculate equilibrium zone profits for Lowe's at \$8,301,573. For Home Depot, we obtain observed and equilibrium zone price profits of \$20,363,560 and \$20,736,293, respectively. Other selection mechanisms, such as the sum of total profits, yields unrealistic equilibrium profits given observed sales.

If both firms choose zone pricing in all markets, in all periods, we must solve for nearly 7,000 prices. The other three possible outcomes have even more prices to solve for. The optimality conditions on firms' problems are highly nonlinear, so we solve for equilibria using state of the art solvers. To search for equilibria, we set 1,000 random starts and solve for the fixed point. On average, around one-third of the starts converge to a fixed point – a BNE for the regime choice.

Since both firms utilize zone pricing, we solve for two parameters of the meta game: a lower point on the menu costs associated with adopted market level pricing. Table 2.8 provides the payoff matrix associated with the metagame. In the absence of menu costs, both Lowe's and Home Depot would adopt market level pricing; however, in this case, Lowe's obtains lower profits with market level pricing than with zone pricing. This is due to both higher demand for Home Depot products in general, as well as a cost advantage in the stores located around Salt Lake City, where Home Depot has a distribution center, but Lowe's does not. This allows Home Depot to undercut Lowe's and gain market share. This analysis shows with a finer degree of pricing – in this case the ability to discriminate at the market level – does not lead to larger profits, a possibility noted in Holmes (1989). Indeed, additional competition hurts Lowe's, but provides a nearly 14% increase in profits for Home Depot.

Also in the absence of menu costs, we find that moving to market level pricing for a single firm increases profits for that firm. Lowe's sees a 4.8% increase in pre-menu cost profits by moving to market level pricing with Home Depot keeping its zone structure. Home Depot also sees modest gains with this regime at 0.97%. On the other hand, if Home Depot moves to market level pricing but Lowe's keeps its zone structure, Home Depot sees a 4.8% increase in profits, largely due to the ability to discriminate in monopoly markets. However, with this regime, Home Depot's competitive advantages, both in costs and demand, results in a 5.9% decrease in profits for Lowe's.

For zone pricing to be the solution of the metagame in pure strategies, it must be the case that $\mu^L \geq \$404,172$ and $\mu^{HD} \geq 2,694,744$. These numbers represent lower bounds on the menu costs associated with adopting market level pricing. These equate to 4.8% and 13.0% of profits for Lowe's and Home Depot, respectively. Together, this yields a menu cost of 10.3% of industry profits, about half the figure calculated using single product deviations. Other market level pricing equilibria exist, some giving much higher profits that match or exceed the menu costs of Section 5. The lower profits in the selected equilibrium could reflect the substitutability between products. Price decreases on one product could prompt the rival firm to discount other products in the same way that it lowers price on the initial product. The increased competition on all products (instead of on only one product) may reduce market share, prompting further rounds of discounting, and lower profits.

Table 2.9 provides summary statistics at the market level across the various pricing regimes of the metagame. The table also provides a summary of the equilibrium in which firms use uniform pricing. We find Lowe's would earn higher profits under uniform pricing than zone pricing (utilizing our selection mechanism). This is consistent with Lowe's earning lower profits under market level pricing than zone pricing; that is, Lowe's benefits when Home Depot has limited ability to price discriminate. Under zone pricing, Home Depot balances the benefits of discriminating in monopoly markets with its desire to undercut Lowe's in duopoly markets. This allows Lowe's to capture market share in duopoly markets than it would not if Home Depot priced at the zone or market level. With uniform pricing, Home Depot obtains approximately \$18.8 million in profits, \$1.8 million less than when both firms use zone pricing and nearly \$5.0 million less than when both firms use market level pricing.

Since Home Depot operates several monopoly stores as part of larger zones, finer pricing results in monopoly prices in these markets, whereas with uniform and zone level pricing, Home Depot balances discriminating in these markets with competing with Lowe's in other markets. The relationship between zone structure and profits is opposite for Lowe's. With a cost disadvantage in the large Salt Lake City market, lower mean utility overall for products, and few monopoly stores, Lowe's does not capture additional profits from finer pricing. Instead, the chain benefits when Home Depot has reduced ability discriminate in competitive markets.

Table 2.8: Metagame of market or current zone pricing

Lowe's / Home Depot	Zone Level Pricing	Market Level Pricing
Zone Level Pricing	\$8,301,573, \$20,736,293	\$7,681,509 , \$21,667,885$-\mu^{HD}$
Market Level Pricing	\$8,705,745$-\mu^{L}$, \$20,937,713	\$7,818,741$-\mu_{L}$, \$23,632,457$-\mu^{HD}$

The zone numbers are equilibrium zone profits instead of observed profits.

Table 2.9: Chain Performance Across Pricing Regimes, summarized across markets

		(1) Zone BNE	(2) Market BNE	(3) Uniform BNE	(4) HD Dev. BNE	(5) Lowe's Dev. BNE
Lowe's						
	mean	276,719	260,625	290,048	256,050	290,192
	median	135,203	120,247	132,386	120,634	124,887
	lower quartile	51,632	57,659	61,037	63,012	64,053
	upper quartile	333,312	283,439	349,276	277,780	270,362
	min	39,470	26,298	50,449	42,432	25,423
	max	1,650,275	1,496,118	1,777,272	1,500,762	1,714,356
	total profits	8,301,573	7,818,741	8,701,427	7,681,509	8,705,745
Home Depot						
	mean	441,198	508,721	401,253	461,019	445,483
	median	243,045	274,062	176,794	262,462	241,781
	lower quartile	142,205	196,623	125,986	191,048	149,604
	upper quartile	435,670	496,624	382,818	418,650	438,875
	min	50,292	23,443	55,720	49,148	71,889
	max	2,979,718	2,906,893	2,868,559	2,842,305	2,916,245
	total profits	20,736,293	23,909,881	18,858,886	21,667,885	20,937,713

Results are annualized and aggregated to the market level.

Chapter 3

Migration and Agglomeration Among Motor Vehicle Parts Suppliers

3.1 Introduction

Motor vehicle manufacturing in North America has been migrating steadily for the past thirty years. Some policymakers have offered large subsidies to new assembly plants, hoping not only to lure the assembly plant but also to influence the migration of parts suppliers that supply the assembly plants. The effectiveness of such a strategy depends on how closely suppliers follow the assembly plants, how pronounced the agglomeration benefits of being near other suppliers are, and what the other factors determining supplier profitability are.

Because of their economic importance and policy relevance, industrial location decisions have inspired a sizable literature both in the motor vehicle parts industry and generally. Many studies have followed Carlton (1979) in applying a discrete choice framework to the site selection of new plants that had entered over a single set period. Early analysis of parts supplier location decisions, beginning with Smith and Florida (1994), followed such an approach. More recent work has included sophisticated variations of discrete choice modeling, but all have treated part supplier entry decisions as static decisions based only on conditions at the time of entry.[1] When plants are durable and require sunk construction costs, the profitability of a location will depend on how location characteristics evolve over the plant's life. Site selection decisions therefore must involve dynamic considerations.

If entrants expected local conditions to remain constant from period to period, static and dynamic models would give the same results. In motor vehicle manufacturing, conditions have been anything but constant. The sustained migration of assemblers forces suppliers to consider dynamics. An entrant valuing proximity to assemblers will evaluate regions with small but growing assembly clusters differently than regions without growth in assembly. Likewise, such a supplier would be more hesitant to enter a region where assembly plants are exiting or contracting. The steady growth of assembly in southern states like Alabama and the steady declines in Michigan over the last 25 years prevent suppliers from relying on conditions at the time as a proxy for the conditions they should expect over the life of a plant. While the migration complicates the problem for both suppliers and econometricians, it also provides the variation in the data needed to separate the effects of observed and unobserved location characteristics and produce unbiased estimates.

One reason dynamics have been omitted in previous analyses of location decisions is they are technically challenging to include. This analysis has several relevant variables in hundreds of locations, creating a state space that until recently would have been intractably large. New two-step techniques for estimating industry dynamics[2] allow such models to be estimated. This paper adopts these techniques to estimate a dynamic entry and exit model where new parts supplier entrants choose between all locations. The first stage of estimations finds policy rules for entry and exit, as well as the transitions of location characteristics, directly from data. These recovered policy rules and transitions allow conditions in a location and their associated payoffs to be simulated. The average discounted lifetime sums of these variables enter the value function. Using these simulated values, the second stage estimates structural value function parameters

---

[1] See also Woodward (1992) and Seim (2006). Static models of motor vehicle parts suppliers entry include Smith and Florida (1994), Klier and McMillen (2008), and Rosenbaum (2012). Head, Ries, and Swenson (1995) add state time trends to an otherwise static model.

[2] Bajari, Benkard, and Levin (2007),Aguirregabiria and Mira (2007), and Pakes, Ostrovsky, and Berry (2007); applications include Collard-Wexler (2010) and Ryan (2012)

that rationalize the observed behavior, as estimated in the first stage. This generalizes the approach of earlier works, with their results corresponding to the first stage policy estimates. The structural parameters recovered in the second stage permit policy simulations that use expectations and replicate the suppliers' dynamic decision.

The model finds that proximity to assembly plants increases the profits of suppliers, but by a small enough amount that single assembly plants have a negligible effect on supplier location decisions. This finding is driven by the small response of observed entry to the enormous shift in assembly locations in the last 25 years. Entry persists in Michigan and the Great Lakes region despite sustained reductions in assembly output, and the quadrupling of assembly in the South has brought a much less pronounced migration of suppliers. This result is found in the first stage policy function estimates as well as the second stage dynamic estimates. The finding is present also in the static entry models similar to the first stage estimation (Klier and McMillen 2008). The most notable difference between the static and dynamic results are with the profitability from locating near other parts suppliers. Head, Ries, and Swenson (1995) first noted the presence of these agglomeration benefits, but static estimates underestimate their size. The relatively stable supplier base continues to attract new entrants to the North, even though assemblers are exiting the region.

The entry and exit model allows policy simulations that move assembly plants, which is effectively what state and local governments do with subsidy bids. The variables of interest, the number of suppliers in each location, are simulated by allowing entrants and exits to use their policy rules in a dynamic game. (See Benkard, Bodoh-Creed, and Lazarev (2010) for a methodologically similar counterfactual experiment.) The counterfactual experiment in section 8 predicts the evolution of supplier locations if the assembly plant recently opened in Tennessee had instead been placed in Michigan. For comparison, I also run model simulations with the actual placement of supplier plants. The supplier counts in the two simulations diverge from each other, responding both to different assembly plant locations and to the endogenous supplier counts themselves. But the two simulations do not diverge much. Since the effect of assembly plant proximity on supplier profits is so small, moving the assembly plants produces a cumulative effect of only one less supplier after four five-year periods. A separate experiment restricts the ability of a entrant to receive agglomeration benefits, and it finds such an entrant has a probability of entry into Michigan that is only a quarter of an entrant with using the actual profit function estimated.

The potential bias caused by unobservable location characteristics confounding the effects of variables of interest had not been directly addressed for part suppliers. Greenstone, Hornbeck, and Moretti (2010) in looking for productivity spillovers from large, new manufacturing plants were concerned that the unknown factors driving those plants to enter could also drive productivity or entry in other plants nearby. In this application, unobservable location characteristics that contribute to supplier entry may easily have influenced assembler or supplier entry in previous periods and therefore be correlated with assembler proximity and incumbent counts. While the approach here, using location fixed effects, differs from the regression discontinuity design of Greenstone, Hornbeck, and Moretti (2010), it addresses the same endogeneity concerns.

Location characteristics determine the profits that drive closure and site selection decisions in the model. The location characteristics considered here correspond to the reasons for clustering advanced by urban economic theory: proximity to customers is measured with proximity to assembly production, peer agglomeration effects are permitted with the number of other suppliers entering profit functions, and cost variables are included to account for the natural advantages of locations. The results contribute to the literature using industry migration to examine the causes of clustering, as in Holmes (1999) and Dumais, Ellison, and Glaeser (2002), and more generally to the literature testing which theorized causes of clustering are empirically important as in Ellison, Glaeser, and Kerr (2010).

The use of a dynamic entry and exit model to estimate agglomeration benefits is similar to Brinkman, Coen-Pirani, and Sieg (2012), but here entrants choose from a much wider set of locations. The two-step estimation techniques allows this paper to recover parameters without assuming an equilibrium selection rule or even calculating equilibrium. Unlike Brinkman, Coen-Pirani, and Sieg (2012), this paper does not estimate productivity or use firm age and size. This model is driven entirely by location characteristics, entry counts, and exit probabilities.

3.2 Industry Migration

Motor vehicle manufacturing in North America is geographically clustered and has been for nearly the entirety of the industry's history. Auto alley, defined here as Wisconsin, Michigan, Illinois, Indiana, Ohio, Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Alabama, and Mississippi, hosts 82% of the motor vehicle final assembly and 76% of the employment at parts suppliers for the United States. Northern Auto Alley, the portion of auto alley north of 40.5° [3], alone contains 41% of national supplier employment and 36% of assembly production.

Suppliers locate near each other and near assemblers. Ellison and Glaeser (1997) found that clustering among suppliers (in SIC 3714) was far more concentrated than if plants were placed randomly according to the population distribution, and that the coagglomeration with assemblers (in SIC 3711) was one of the most pronounced among all upstream-downstream industry-pairs.

Northern Auto Alley traditionally had an even higher concentration of assemblers. In 1986 Northern Auto Alley produced 5.8 million vehicles, over half of the United States total. While the production totals fluctuated with the business cycle, the share of national production by Northern Auto Alley steadily declined. Ford and General Motors had established branch assembly plants near major markets throughout North America in the early 20th century. Starting in the 1980s, they gradually closed their outlying assembly plants, so production in the United States outside of auto alley declined even more than in Northern Auto Alley.

Meanwhile, automakers headquartered in Asia and Europe built "transplant" assembly plants in the Midwest and South. Honda opened an assembly plant in central Ohio in 1982, and Nissan entered in Tennessee in 1984. Toyota began joint operation of a California plant with General Motors in 1983 and built its own plant in Kentucky in 1987. In 2011, Asian and European firms operated 17 assembly plants in the United

---

[3] Michigan, Wisconsin, and the northern thirds of Illinois, Indiana, and Ohio are included by this definition. Figures 1 and 2 display the boundary used on a maps of Auto Alley. Southern Auto Alley is defined as the portion of auto alley south of 40.5°. This cutoff was chosen to group together all the assembly plants wholly owned by foreign carmakers. In estimation, latitude will be discretized in bins, with 40.5° as a divider, but this can be adjusted in robustness checks.

Table 3.1: Migration of motor vehicle manufacturing

	1986	1991	1996	2001	2006	2010
Assembly production (% of US total)						
Northern Auto Alley	51.6	48.8	45.3	41.2	36.7	36.3
Southern Auto Alley	12.1	21.9	26.5	31.4	40.5	45.3
US outside Auto Alley	36.2	29.3	28.3	27.4	22.8	18.4
Parts supplier employment (% of US total)						
Northern Auto Alley	50.3	48.7	46.1	42.6	43.1	41.3
Southern Auto Alley	27.0	26.5	29.1	32.7	32.7	35.1
US outside Auto Alley	22.7	24.8	24.8	24.6	24.2	23.6

States,[4] of which 13 are in the Southern Auto Alley. Because of transplant construction (and despite Ford and General Motors plant closures in Atlanta), the number of assembly plants in the Southern Auto Alley has more than doubled since the early 1980s. Southern Auto Alley eclipsed Northern Auto Alley production counts by 2006, reversing a gap of four million vehicles twenty years earlier. The top half of table 3.2 reports the percent of national assembly output by region and shows the steady migration throughout the 25 year period.

The migration of assemblers was identified while still underway(Rubenstein 1992). Books describing the advantages that transplants entering the south had over the incumbent assemblers concentrated in the north had even reached the popular press.[5] Suppliers of the period would have been aware of the migration and would have needed to consider it when making plans.

The changing landscape of motor vehicle assembly has provided states many opportunities to bid on new plants or compete to retain existing plants. State and local governments offer to pay training costs, build infrastructure, and provide other indirect subsidies for new assembly plants. Every winning bid in the past decade has had a

---

[4] The 2011 count includes Mazda's joint venture with Ford. Ontario, Canada also hosts 1 joint venture and 3 transplant assembly plants.

[5] James Womack, Daniel Jones, and Daniel Roos published *The Machine That Changed the World* in 1990. The first edition of David Halberstam's *The Reckoning* was printed in 1986.

reported value exceeding $100 million; the most recent assembly plant announced for North America, Volkswagen's Chattanooga assembly, followed a subsidy bid reportedly costing $588 million. While assembly plants are large employers, the subsidies governments pay exceed several years of plant payroll. Policymakers argue that a winning assembly plant bid will have a broader economic impact and that parts suppliers will follow the assembly plant. For example, following the Volkswagen announcement, the local press reported Tennessee's "Governor Bredesen said the 2,000 direct jobs at VW are 'the tip of the iceberg.'" Mississippi Governor Haley Barbour said: "We expect more suppliers to begin work and create jobs as the [Toyota Blue Springs] auto plant prepares to start production."[6]

The overall migration of suppliers has been less dramatic than the movement of assemblers, but still notable. Although assembly production between 1986 to 2011 decreased 58% in Northern Auto Alley, employment at supplier plants decreased only 12 % and plant counts increased by 15%. Assembly production during the same period increased in Southern Auto Alley 124%, but supplier employment grew by a more modest 40% and supplier plant counts increased by 73%. The bottom half of table 3.2 shows supplier plant employment in each region.

3.3 Data

I construct a panel of supplier plants and location characteristics for five five-year periods beginning in 1986 and ending in 2011. The data set is limited to 12 eastern states in Auto Alley. Part suppliers outside of auto alley may focus on the aftermarket or export market and have a wholly different set of location preferences from those operating in Auto Alley or considering entry into Auto Alley. The states within Auto Alley accounted for more than 75% of national employment in motor vehicle parts manufacturing in every data period.

For estimation, a county is a location. Location characteristics include local manufacturing unionization rates, interstate highway presence (reported as a binary), local

---

[6] Pare, Mike. "Chattanooga 'best fit' for VW, CEO says" *Chattanooga Times Free Press.* 16 Jul 2008. and "Governor Barbour Announces Toyota Supplier Will Restart Operations in Baldwyn." Office of the Governor of Mississippi Press Release. 30 July 2010.

Table 3.2: Summary Statistics for Location Characteristics

Variable	Mean	Std. Dev.	Min	Max
Interstate in county	0.80	0.40	0	1
Manufacturing wage (hundred $/wk)	5.55	1.30	2.37	8.53
Population density (hundred per sq km)	10.83	15.24	0.12	54.87
Incumbent suppliers in county	11.6	15.7	1.0	50.0
Private sector unionization rate	0.195	0.118	0.000	0.594
Assembly quantity within 100km (mil)	0.923	1.305	0	3.971
Assembly quantity within 700km (mil)	5.898	2.729	0	8.334
Locations	1106			

Data for 1986, weighted by incumbent suppliers.

manufacturing wage (reported as a weekly average in hundreds of 1986 dollars), and the production quantity of assembly plants within 100 and 700 kilometers (both reported in millions of vehicles per year). The data appendix lists specific sources for each variable. Table 3.3 displays summary statistics of these location variables.

Supplier plant locations are based on Dun & Bradstreet data. Supplier plants are establishments to which Dun & Bradstreet assigned the primary SIC classification of 3714 Motor Vehicle Parts and Accessories. This Standard Industrial Classification (SIC) code for parts includes steering wheels, transmissions, brake systems, some engine parts, and most other parts. It accounts for almost half of the inputs used by plants in the motor vehicle assembly classification. Separate classification codes were given to Carburetor, piston, piston ring, & valve mfg (3592), Vehicular lighting equipment (3647), and Engine electrical equipment (3694). Plants classified into these codes are not included in descriptive statistics, but have been used in robustness checks of the model. Other components, like seats and glass, were classified into general categories and will not be covered in this analysis.

Table 3.3: Supplier counts

	1986	1991	1996	2001	2006	2011
Northern Auto Alley	467	552	652	671	646	539
N Illinois	72	68	79	69	58	56
N Indiana	45	57	81	77	82	69
Michigan	228	282	324	360	377	297
N Ohio	93	108	124	120	92	77
Wisconsin	29	37	44	45	37	40
Southern Auto Alley	263	361	411	454	496	456
Alabama	18	20	20	23	31	37
Georgia	29	33	29	26	31	39
S Illinois	19	21	18	24	23	20
S Indiana	47	50	61	84	88	71
Kentucky	21	33	47	60	69	60
Mississippi	13	16	16	14	16	16
North Carolina	31	49	57	42	51	46
S Ohio	36	58	63	60	56	48
South Carolina	11	22	32	42	50	51
Tennessee	38	59	68	79	81	68

A dynamic entry and exit model requires a panel data set. The Dun & Bradstreet data was selected because its cross sections can be linked. (For details of that procedure, see Appendix A1.) Entry is defined as the appearance of a plant not open in any previous period, closure as a plant not present in any future period. The Dun & Bradstreet data for this period replicates patterns in plant counts found in County Business Patterns and is suited for entry and exit models. Appendix A2 discusses the quality checks on the Dun & Bradstreet data further.

Table 3.3 shows parts supplier plant counts. Note that in each cross section the plant count and employment totals (shown in table 3.3) are highly correlated, though average employment declines in all regions over time. For simplicity, the model will use plant counts.

Table 3.4: Part supplier employment (thousands)

	1986	1991	1996	2001	2006	2011
Northern Auto Alley	171.3	203.4	199.0	231.1	186.7	151.6
N Illinois	13.7	11.7	13.0	14.2	8.1	8.8
N Indiana	10.0	11.8	17.7	16.0	15.2	15.3
Michigan	96.7	129.0	121.2	145.5	126.5	92.7
N Ohio	39.5	37.5	37.5	43.6	30.6	27.3
Wisconsin	11.3	13.4	9.7	11.7	6.4	7.5
Southern Auto Alley	91.9	110.8	125.3	177.6	141.6	128.9
Alabama	4.2	6.0	8.7	10.4	7.0	6.7
Georgia	3.5	4.0	3.4	6.1	5.4	5.2
S Illinois	4.5	5.9	6.5	8.1	8.7	7.0
S Indiana	33.1	27.1	35.7	55.9	34.5	36.6
Kentucky	4.0	8.2	11.0	16.2	17.1	16.0
Mississippi	3.4	4.2	4.3	4.7	2.5	2.7
North Carolina	6.3	12.7	12.4	12.9	11.3	10.2
S Ohio	17.4	25.8	19.3	31.8	21.7	16.0
South Carolina	2.2	4.0	7.5	11.3	13.3	11.4
Tennessee	13.2	13.1	16.6	20.1	20.1	17.1

Table 3.5: Supplier entry counts

	New entrant count for period beginning in				
	1986	1991	1996	2001	2006
Northern Auto Alley	299	287	258	163	115
Southern Auto Alley	212	174	182	141	100

New entrant counts are shown in table 3.3. Strikingly, the majority of new supplier entrants every period chose to locate in the north despite the dramatic exit of assemblers.

3.4 Model

The entry and exit model focuses on the location decisions of part suppliers. Incumbent supplier plants and potential entrants are players in a dynamic game of the class introduced by Ericson and Pakes (1995). Time is discrete in this model. In each period potential entrants choose if and where to enter. Incumbent supplier plants decide whether to exit the industry or remain in operation. New entrants and remaining incumbents earn profits based on their locations and the locations of other suppliers. After a new entrant selects its location and becomes an incumbent supplier, and its only remaining decision is when to exit.

Location characteristics drive supplier profits in the model. Locations are indexed by ℓ. At any time t, each location has a set of characteristics $X_{\ell t}$, which suppliers take as given. Assemblers are not players in this model, but assembler proximity is a component of $X_{\ell t}$. While supplier locations in the aggregate may influence assembly plant placement, as in Holmes (2004), the impact of any single supplier plant is negligible, so the suppliers modeled treat assembler proximity as exogenous. Modeling assembler and supplier location decisions jointly would increase the complexity, but would not greatly change estimation of the individual supplier decisions that are the focus of this work.

Every period each incumbent plant must decide whether to remain open or close permanently. Suppliers take all location characteristics and competitor locations as given. At the beginning of each period, each supplier draws a private, random scrap value ϵ_{it}^{exit} and a profit shock ϵ_{it}^{stay}. After observing these draws the incumbent may either claim the scrap value by exiting permanently or remain and earn period profits $\pi(X_{\ell t}, n_{\ell t}, \xi_{\ell}) + \epsilon_{it}^{stay}$, where

$$\pi(X_{\ell t}, n_{\ell t}, \xi_{\ell t}) = \gamma_X X_{\ell t} + \gamma_n n_{\ell t} + \xi_{\ell t}$$

All production costs enter period profits linearly, that is as fixed costs, so that in estimation the parametric functional forms will be as simple and transparent as possible. Profits also depend on unobservable characteristics, $\xi_{\ell t}$ for each location. The unobservable characteristics may be correlated with $n_{\ell t}$ and $X_{\ell t}$, which would bias estimates that ignore $\xi_{\ell t}$. The correlation with the number of incumbents $n_{\ell t}$ arises because the unobservable characteristics suppliers find profitable would have prompted more suppliers to enter in previous periods. Assembler proximity is a component of $X_{\ell t}$ and the unobservable characteristics that benefit suppliers may also attract assembly plants. Greenstone, Hornbeck, and Moretti (2010) were concerned about bias caused from local unobservables enough to use a regression discontinuity design in their productivity estimation. This paper will offer several approaches to control for unobservable characteristics, most of them made possible by panel data and variation over time brought by the migration.

The scrap value is drawn from a Type I extreme value distribution. The profit shock is the sum of two random variables. The first is a private, idiosyncratic component drawn from a Type I extreme value distribution, the second a time-specific shock common to all suppliers. Motor vehicle production and profitability are highly cyclical, but the focus of this model is on supplier location choice and not the determinants of the macroeconomy. Some of the effects of recessions will come through in the wages and assembler production quantities contained in X_t, but all other macroeconomic conditions relevant to the closure decision of a supplier plant will enter into ϵ_{it}^{stay}.

Suppliers at the same location differ only in the idiosyncratic profit shocks and scrap values they draw. Since this model of location choice compares costs among locations without plants, plant-specific productivity measures are not estimated. (For production function estimation among motor vehicle assembly plants, see Van Biesebroeck (2003).)

A plant that remains open may operate or close and claim a scrap value in a future period, so its decision is a dynamic one. Each incumbent has expectations over the state variables $(X_{1t}, X_{2t}, \ldots, X_{Lt}, n_{1t}, n_{2t}, \ldots, n_{Lt})$. (For notational compactness, let

$X_t = \times_{\ell=1}^{L} X_{\ell t}$, $n_t = \times_{\ell=1}^{L} n_{\ell t}$, and $xi_t = \times_{\ell=1}^{L} \xi_{\ell t}$). Let $a_{it} = 1$ denote closure for incumbent i in period t. Then the incumbent value function is

$$V_\ell(X_t, n_t, \xi, \epsilon_{it}^{inc}) = \max_{a_{it}} \begin{cases} \pi(X_{\ell t}, n_{\ell t}, \xi_{\ell t}) + \beta E_\ell[V_\ell(X_{t+1}, n_{t+1}, \xi_{t+1}, \epsilon_{it+1})|X_t, n_t] + \epsilon_{it}^{stay} & : a_{it} = 0 \\ \epsilon_{it}^{exit} & : a_{it} = 1 \end{cases} \quad (3.1)$$

Expectations are indexed by ℓ as state variables can evolve according to different processes in different locations. The value function is indexed by ℓ only because expectations are also. Let χ_ℓ be the exit policy function that solves the value function in location ℓ.

Potential entrants select their location at the same time as incumbents make exit decisions. At the beginning of each period potential entrants draw location-specific entry costs for each location. They decide if and where to enter based on each location's incumbent value and the entry cost drawn for it.

Like the scrap value, the entry cost that potential entrant k draws for location ℓ has two components: $\epsilon_{k\ell t}^{entry} = \phi_t + \kappa_{k\ell t}$. The first component, ϕ_t, is a period-specific common cost drawn each period. As with exit, the relative attractiveness of entering or remaining outside the industry fluctuates with the macroeconomy. The second component, $\kappa_{k\ell t}$, is drawn from a Type I extreme value distribution and represents the entrant-specific considerations in the site selection problem.

All new entrants simultaneously choose a location in which to enter and become an incumbent or choose an outside option and remain outside the industry forever. The selection rule will be

$$\mu(X_t, n_t, \xi, \epsilon_t^{entry}) = \arg\max_l \begin{cases} \epsilon_{k0}^{entry} & : l = 0 \\ E[V_l(X_t, n_t, \xi, \epsilon_{kt})] - \epsilon_{kl}^{entry} & : l \in \{1, 2, \ldots, L\} \end{cases} \quad (3.2)$$

where $l = 0$ is the action for not entering and ϵ_{k0}^{entry} the random value of the outside opportunity. The entrants minimize their value minus entry costs through their site selection. Let $\mu(X_t, n_t, \xi, \epsilon_k^{entry})$ be the corresponding selection rule.

In this model an equilibrium is a policy function for potential entrants $\mu(X, n, \xi, \epsilon_k^{entry})$, a policy function for incumbent suppliers in each location $\chi_\ell(X_\ell, n_\ell, \xi, \epsilon_i^{inc})$, transition probabilities $g_\ell(X_\ell', X_\ell)$, and a value function for each location $V_\ell(X, n, \xi, \epsilon_i^{inc})$ such that (1) given (X, n, ϵ, g), χ_l maximizes the value function in equation 1, (2) given

$(X, n, \xi, \epsilon_k^{entry})$, $\mu(X, n, \xi, \epsilon_k^{entry})$ solves the entrants' maximization problem in equation 2, (3) supplier counts evolve according to $n'_\ell = n_\ell + \sum_k I\left(\mu(X, n, \xi, \epsilon_i^{entry}) = \ell\right) - \sum_i^{n_\ell} \chi_\ell(X, n, \xi, \epsilon^{inc})$, (4) expectations on n_ℓ are rational, and (5) expected transition of exogenous state variables $g_\ell(X'_\ell, X_\ell)$ match those observed.

3.5 Estimation

3.5.1 Specification

For estimation, locations will be the 550 counties in auto alley that hosted a supplier plant sometime since 1986. The location characteristics, $X_{\ell t}$, include: interstate presence (interstate), local manufacturing wage (wage), state manufacturing unionization rate (union), and population density (popdenisty), quantity at assembly plants within 100 kilometers (qw100km), and quantity at assembly plants within 700 kilometers (qw700km). Interstate presence was largely constant throughout the 25 year data window and so is treated as a static variable. Unionization rates are a static variable, since the greater geographic detail needed would not have been available period by period. The supplier count n and all other components of X will be treated as dynamic state variables.

Only suppliers in the same county enter $n_{\ell t}$ and therefore period profits. The most widely proposed agglomeration mechanisms, such as labor pooling and knowledge spillovers, work on a very local level. Congestion costs related to overcrowded local infrastructure also are largely contained within a county. The influence of assembly plants is spread more widely, hence the assembler proximity measure in $X_{\ell t}$ account for plants far outside county boundaries. This unfortunately makes comparisons of assembly plant and supplier plant influence more difficult.

The unobserved characteristics $\xi_{\ell t}$ will be assumed to be permenant, and therefore location fixed effects can be used to recover the effects on entry and exit. Persistant characteristics are the ones most likely to be correlated with the entry decisions of assemblers and other suppliers that occurred in previous periods and therefore to the assembly proximity measures and supplier counts in the current period.

Periods will be five years and locations will be counties in auto alley. The number of potential entrants each period will be set at 700. The discount factor suppliers use will be $\beta = 0.90^5 \approx 0.59$, since periods are five years long. Robustness checks with different discount factors and potential entrant counts produce similar results.

3.5.2 Procedure

Estimation proceeds in two stages, following Bajari, Benkard, and Levin (2007). In the first stage, the entrants' location selection policy rule μ and incumbents' exit policy rule χ are taken directly from the data using flexible estimation. Transition functions for the exogenous state variables also are estimated. Using these estimates, forward simulations are run for each location and each period to find the discounted lifetime profits in terms of the model parameters. Finally, the second stage finds the model parameter values that best fit the observed behavior of potential entrants.

Locations are classified into two discrete bins by latitude, using the 40.5° latitude cutoff introduced in section 2. Each dynamic state variable in each latitude bin is assumed to follow first-degree autoregressive processes dependent only on itself. Separate regressions for each latitude bin are employed because different processes seem to be at work in the north and south during the migration.

The exit policy function $\chi(X_t, n_t, \epsilon_i)$ is estimated semiparametrically, with a flexible function of local state variables added to a logit error term. High order and interaction terms are included for the most important variables, while, for now, the characteristics of other locations are omitted from the regression.

To recover the entrants' selection rule, the expected incumbent value at each location is approximated by function of local state variables with higher order terms and interactions, but the estimation assumes the selection rule follows the multinomial discrete choice structure presented in the model. This multinomial logit accounts for the state variables for all locations, albeit in specific parametric form.

The existing static entry models for this industry follow a procedure similar to this policy function estimation. Indeed, one way to interpret the results of previous studies is as the policy function for far-sighted entrants.

Year fixed effects are included in the estimation of both the entry and exit policy functions to account for the common year-specific components in entry costs and profit shocks. Location fixed effects can control for the unobservable characteristic ξ_ℓ and prevent the endogenity bias that would otherwise occur.

Using the decision rule estimates and the observed transition frequency, the total expected lifetime discount profits of an incumbent are computed for each county-period observation. State variables and the exit decisions are forward simulated and the discounted sum for each variable recorded. The average of 500 simulations for each observation is used.

The second stage finds what parameters best fit the observed behavior of potential entrants. Instead of the minimum distance estimators originally proposed in Bajari, Benkard, and Levin (2007), estimation will again use the multinomial choice structure of entry decision. The Berry inversion allows for the difference in log shares of entrants in each location and the log shares of potential entrants remaining outside the industry to be regressed on the total lifetime discounted sum for each variable an incumbent expects in each location.

3.6 First Stage Results

First stage estimates of entry deliver the relative probabilities of entry in each location expected by incumbents. Table 3.6 reports five specifications. The first four columns contain only first order terms for ease of interpretation. They differ only by the type of location fixed effect used to control for the persistent unobservable location characteristics ξ. The unobservables are potentially correlated with assembly proximity and the number of incumbent suppliers, but the inclusion of location indicators affects coefficients on the variables of concern only minimally.

The coefficient signs largely match those found in established literature for this industry with suppliers seeking locations with interstates, lower population densities, and low union activity. The presence of additional assembly plants increases entry probability, something not clear in Klier and McMillen (2008).

Estimates for the autoregressive processes governing the exogenous state variables are reported in table 3.6. Separate processes are assumed to operate in each latitude bin, and indeed parameters for each group differ. The migration of assembly plants into the south means that a location in the south should expect more assembly within

Table 3.6: First stage entry policy function estimation

	(1)	(2)	(3)	(4)	(5)
suppliers	0.0521***	0.0521***	0.0504***	0.0551***	0.115***
	(0.00220)	(0.00222)	(0.00231)	(0.00660)	(0.0110)
qw100km	0.0946***	0.0945***	0.0858**	0.336**	0.0210
	(0.0273)	(0.0273)	(0.0287)	(0.104)	(0.134)
qw700km	0.00423	0.00409	0.0237*	0.0453	0.00730
	(0.00492)	(0.00510)	(0.0115)	(0.0306)	(0.0192)
interstate	0.0630*	0.0630*	0.0576*	.	0.00137
	(0.0280)	(0.0280)	(0.0285)	.	(0.0265)
popdensity	0.0118***	0.0118***	0.0138***	-0.0506	0.0185***
	(0.00230)	(0.00231)	(0.00244)	(0.0315)	(0.00501)
union	-0.0597	-0.0650	0.00836	.	0.0533
	(0.135)	(0.145)	(0.167)	.	(0.311)
wage	0.0257	0.0257	0.0269*	-0.0167	0.0290
	(0.0132)	(0.0132)	(0.0136)	(0.0536)	(0.0574)
wage^2					-0.00269
					(0.00563)
union^2					0.491
					(0.634)
popdensity^2					-0.000453***
					(0.000116)
suppliers^2					-0.000920***
					(0.0000958)
qw100km^2					0.0290
					(0.0245)
qw700km^2					-0.000948
					(0.00159)
qw100km × qw700km					-0.00305
					(0.0154)
qw100km × suppliers					-0.00365
					(0.00260)
qw700km × suppliers					0.000386
					(0.00115)
Location indicators	none	latitude bin	state	county	latitude bin
Year indicators	yes	yes	yes	yes	yes

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.7: Transition of local characteristics

	Lagged variable coeff	Constant	Sq Error
	Southern Auto Alley		
Population density	1.0336	0.0577	0.3669
Wage	0.8925	0.5173	0.4692
Assembly quantity with 100 km	0.8775	0.0322	0.1277
Assembly quantity with 700 km	0.8649	0.5870	1.2381
	Northern Auto Alley		
Population density	0.9919	0.1489	0.5152
Wage	0.8345	0.7897	0.5027
Assembly quantity with 100 km	0.8646	0.0238	0.2584
Assembly quantity with 700 km	0.6009	3.1980	1.3489

100 kilometers next period than should a northern location with the same quantity this period.

Logit regressions describing the incumbent exit rule are reported in table 3.6. Again the simpler first specification is only to aid interpretation. The forward simulations use the second specification. Suppliers are more likely to exit if operating in a location with high wages, high union membership rates, and with little assembly within 100 kilometers. Note that the relative importance and even direction of some variables differ in the estimated entry and exit policy functions. Different plant turnover rates in the data drive these differences.

3.7 Second Stage Results

The second stage coefficient estimates for the full model are reported in table 3.7. Supplier profits are decreasing in local wages and unionization rates. The coefficient for suppliers, γ_n in the same county is positive, indicating that agglomeration benefits outweigh any costs of local competition. The model cannot distinguish between different sources for the agglomeration benefits, but benefits measured are in addition to the

Table 3.8: First stage exit policy function

	(1)	(2)
suppliers	-0.00426	0.0504
	(0.00278)	(0.0303)
qw100km	0.0918	-0.0772
	(0.0533)	(0.408)
qw700km	-0.0681***	-0.292***
	(0.0150)	(0.0730)
wage	0.0831*	-0.114
	(0.0397)	(0.214)
union	0.460	1.531
	(0.356)	(0.934)
interstate	-0.0371	-0.0350
	(0.0886)	(0.0938)
popdensity	0.00480	0.0239*
	(0.00334)	(0.0118)
suppliers^2		-0.000101
		(0.000122)
qw100km^2		0.0655
		(0.0517)
qw700km^2		0.0183**
		(0.00565)
qw100km × qw700km		-0.00328
		(0.0424)
qw100km × suppliers		-0.00339
		(0.00341)
qw700km × suppliers		-0.00432
		(0.00288)
wage^2		0.0151
		(0.0203)
union^2		-1.565
		(1.765)
popdensity^2		-0.000441
		(0.000225)
northern auto alley	0.334***	0.319***
	(0.0852)	(0.0896)
Year indicators	yes	yes
N	4973	4973

natural advantages of a location (measured with local characteristics in X_ℓ and ξ_ℓ) or the need to be near the final customers, the assembly plants. Since many parts suppliers have other parts suppliers as their customers in a multi-tiered network, some of the agglomeration may reflect the transportation costs associated with receiving or sending intermediate goods. Nearby assembly plants increase profitability, but only slightly.

For comparison, dynamic results are compared to estimates of the static model. Suppliers in the dynamic model consider being in a region with high assembly quantity to be less important. The static model explains the persistence of suppliers in Northern Auto Alley by noting the high contemporaneous assembly production. In the dynamic model, suppliers see the transition of assembly quantity within 700 kilometers and know that high current assembly production is no guarantee of continued assembly production. Since suppliers persist in the north even though they suspect assemblers will leave, the dynamic model must find other, less transitory factors to explain the continued entry in the north. The dynamic model concludes that agglomeration effects must be bigger and that union costs must be smaller than static models would suggest.

Assembly quantity within 100 kilometers also is more important in the dynamic model than in the static estimates. Areas with the highest concentration of assembly plants also have the highest turnover rates and therefore the highest predicted exit rates and shortest expected plant lifetime. Yet entrants still select these areas. The model concludes that entrants must find the immediate proximity of assembly plants profitable enough to justify the higher hazard rate caused by other variables.

Table 3.7 reports the marginal effect of a unit increase of each variable on the expected number of new entrants in 2006 in a location with average state variables. The dynamic model differs from the static model not just in coefficient values, but in that future state variable values are used to predict entry. The table calculates the effect of both a single period increase in each variable while holding the rest of a supplier's profit stream fixed and a permanent proportional increase in the variable in all future periods. A one million vehicle increase in nearby assembly production, the rough equivalent of three new assembly plants, brings up the number of supplier plants entering each county by only a small fraction.

Table 3.9: Dynamic and myopic results

	Static (1)	Dynamic (2)
suppliers	0.0521***	0.192***
	(0.00222)	(0.00802)
qw100km	0.0945***	0.318***
	(0.0273)	(0.0831)
qw700km	0.00409	-0.0109
	(0.00510)	(0.0144)
wage	0.0257	0.00431
	(0.0132)	(0.0474)
union	-0.0650	-0.0744
	(0.145)	(0.397)
interstate	0.0630*	0.117
	(0.0280)	(0.0709)
popdensity	0.0118***	0.0351***
	(0.00231)	(0.00704)
Year indicators	Yes	Yes
Location indicators	latitude bin	latitude bin

Standard errors in parentheses. Standard errors currently account only for second stage and therefore are a lower bound.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.10: Marginal effect on the number of new entrants

	Static model	Dynamic model single period change	Dynamic model permanent change
suppliers	0.022	0.054	0.093
qw100km	0.040	0.094	0.167
qw700km	0.002	-0.002	-0.005
wage	0.011	0.001	0.002
union	-0.025	-0.019	-0.031
interstate	0.026	0.033	0.054
popdensity	0.005	0.010	0.016

Marginal effects at the average values, measured by number of new entrants per period.

3.8 Counterfactuals

3.8.1 Simulation with counterfactual placement of assembly plants

The dynamic model can simulate supplier entry and exit starting from actual or counterfactual conditions. The transition functions estimated in the first stage can be used to simulate location characteristics in future periods, in much the same way that they were used in the forward simulations in the estimation routine. Because the policy functions estimated in the first stage are equilibrium objects, they too can be used in the forward simulation to simulate the site selection of new entrants and the closure decisions of incumbents.

The model can be used for the exact counterfactual experiments policymakers should consider. The additional number of suppliers brought to a jurisdiction by a successful bid can be estimated by comparing model simulations where assembler proximity variables are increased to reflect the new assembly plant locating in different sites.

The experiment here specifically moves the new Volkswagen assembly plant in Chattanooga, Tennessee to Grand Rapids, Michigan. The counterfactual location was motivated by press reports that before announcing their location decision, Volkswagen offficals had considered three finalist sites in Alabama, Tennessee, and Michigan. Tennessee's efforts to become the host of the new plant included a subsidy bid reportedly costing \$577 million.[7] If a smaller subsidy offer from Tennessee would have lead Volkswagen to place its plant in Michigan, this counterfactual experiment answers how many suppliers Tennessee's half billion dollar subsidy brought.

In the first model simulation all assembly plants are given their actual locations and actual production quantities for 2011. (Since the new Volkswagen plant opened near the end of 2011, I use an estimate of 150,000 units for its production count. This should form a better basis for supplier expectations than the actual count from the last few months of 2011.) From this, the assembly production within 100 kilometers and within 700 kilometers variables are calculated. All other location characteristics start with their 2011 values. Following Benkard, Bodoh-Creed, and Lazarev (2010), the first stage

---

[7] Pare, Mike. "VW Spends Most of \$235 million in infrastructure aid." Chattanooga Times Free Press. 12 Oct 2011. The bid included \$130 million in waived taxes that would not have existed without the plant anyway, but the rest of the subsidies represent real expenditures.

Table 3.11: Supplier plant counts in the counterfactual simulation

	Supplier counts	
Period	Model	Counterfactual
Tennessee		
0 (2011)	68.00	68.00
1 (2016)	74.64	74.27
2 (2021)	78.52	78.02
3 (2026)	80.57	80.14

estimates of transitions and policy function simulate the evolution in each county of the number of suppliers and of all location characteristics, including assembly production nearby. The counterfactual simulation moves the coordinates for the Volkswagen plant to those for Grand Rapids, Michigan. This results in starting period values for assembly production within 100 kilometers and 700 kilometers that are lower for counties near Chattanooga and higher for counties near Grand Rapids. Otherwise, the procedure is indentical to the model simulation. Both model and counterfactual simulation were run 10,000 times; the difference in average supplier counts for each county reflects the influence of the moved plant.

Table 3.8.1 reports the average supplier counts in the model and counterfactual simulations summed for all counties in Tennessee. Both simulations predict a continued increase in the Tennessee supplier base, though the growth is lower in the counterfactual that removes the Volkswagen assembly plant. The two models diverge only slightly, never by more the 0.5 suppliers (representing about 150 supplier jobs). The county with the largest difference between model and counterfactual simulation, was Hamilton County, the home of the assembly plant. Other nearby counties and counties throughout Tennessee with the largest existing supplier bases were also among the places that benefited the most. A few counties in Tennessee even had slightly more suppliers under the counterfactual, since the new assembly plant made them relatively less attractive than counties nearer Chattanooga, which are close substitutes to them. Nevertheless, almost all counties had fewer suppliers in the counterfactual in most periods, though the effect everywhere was small.

The model on which these simulations are built assumes all parts suppliers value the assembler proximity the same. In reality, different transportation costs for particular parts likely cause some heterogeneity. A subset of parts suppliers with a much higher affinity for nearby assembly plants would be more sensitive to the movement of assembly plants than averages would suggest, so a next step is to test the robustness of these results with different sub-classifications of the supplier parts industry.

3.8.2 Entry patterns without agglomeration

The second stage estimates emphasize the importance of a persistent supplier base to the profits of other part suppliers. To see how these agglomeration benefits are driving entry, suppose the profits of one new potential entrant do not depend on the number of suppliers in its prospective locations. That is, let $\gamma_n = 0$ for this one entrant, preventing it from benefiting from any spillovers or agglomeration benefits in excess of local competition or congestion costs. Such an entrant will not affect the overall equilibrium being played or any suppliers expectation of how the industry will evolve, so the first stage estimates will still represent the equilibrium policy functions for all other suppliers. The one new entrant, however, will have markedly different entry probabilities from the rest of its cohort.

Without net agglomeration benefits, a new potential entrant in 2006 would enter Michigan with 6.9% probabilities, while the dynamic model predicts the average probability of entry in Michigan at 28.2%. (Performing a similiar experiment with a static model yeilds probabilites of 5.9% and 15.9% respectively.) Agglomeration benefits are therefore important to the maintaince of the persistent supplier base in Northern Auto Alley.

Probabilities and predictions under a counterfactual where no supplier can benefit from spillovers cannot be calculated, because such an experiment would change enough entry and exit policy functions. The Bajari, Benkard, and Levin (2007) methodology obtains its tractibility by not solving for equilibrium and by using the data to implicitly solve for the equilibrium selection mechanism being used. Therefore, I am not able to calculate what the new equilibrium policy functions would be and so cannot find the evolutionary paths of location characteristics that suppliers expect. The experiment with one new entrant does hint that the impact would be large and that entirely different equilibria would emerge in the absence of peer agglomeration.

3.9 Conclusion

Durability and entry costs make selecting a site for plants a long term decision. An industry migration increases the importance of dynamic concerns, which static models may miss. In the case of motor vehicle parts suppliers, the static model underestimates the extent of peer agglomeration benefits.

A model that can estimate the entry and closure decisions of parts suppliers can be used to estimate the benefits of attracting assembly plants. It finds Tennessee's expensive subsidies of assembly plants have had little influence on the location decisions of parts suppliers.

3.10 Data Appendix

3.10.1 Data Sources

Supplier plant locations come from the Dun's Metalworking Directory for 1996 and before. For 2001 and latter, I use the Dun & Bradstreet Million Dollar Directory omitting plants with fewer than 20 employees to match the Metalworking Directory's inclusion criteria. Plants are matched through time mostly by DUNS number, a permanent identifier of each plant. Because the DUNS number sometimes changed without reason, plants that had no DUNS number match in the subsequent year were also linked by address. (Cases in which matching addresses lacked street numbers were linked only if the company name remained constant or a corporate merger could be verified.) The Dun & Bradstreet data sometimes contains separate records for divisions within the same plant, so exact address duplicates were merged together.

Wage data is from the Bureau of Labor Statistics's Quarterly Census of Employment and Wages (QCEW). The wages used are the county- and year-specific average weekly wage for manufacturing plants (SIC 31-33). In counties where the manufacturing wage is unavailable, the state average manufacturing wage is used.

Population estimates and county areas are from the US Census. Union membership rates are state level from the Union Membership and Coverage Database. The construction of that database from the Current Population Survey is described in Hirsch and MacPherson (2003).

Interstate highway indicators were constructed from map files published by the National Atlas. Because of the stability in the interstate system since 1986, highway presence is a static variable that uses current data. The location and production quantities for assembly plants are from Ward's Automotive Yearbook. (Pending the release of 2011 production figures, 2010 counts are used in their place.)

3.10.2 Quality of Plant Panel

Some early Dun & Bradstreet data are known to overreport employment, to overreport plant counts, and to detect new entrants belatedly. Neumark, Wall, and Zhang (2011) find that in a data set based on Dun & Bradstreet data from 1992 to 2006 employment measures are higher than in the QCEW or the Current Employment Statistics (CES),

Table 3.12: Supplier counts (alternate data source: CBP plants with 20+ employees)

	1986	1991	1996	2001	2006	2009
Northern Auto Alley	418	486	535	602	619	496
N Illinois	57	65	61	69	65	58
N Indiana	45	67	80	99	97	64
Michigan	208	241	262	275	284	235
N Ohio	81	83	94	107	120	95
Wisconsin	27	30	38	52	53	44
Southern Auto Alley	253	343	419	525	575	504
Alabama	18	18	20	29	46	51
Georgia	19	25	29	36	44	31
S Illinois	11	17	15	18	22	20
S Indiana	34	49	57	77	80	72
Kentucky	15	34	47	69	77	73
Mississippi	16	25	24	30	23	18
North Carolina	36	41	59	66	66	59
S Ohio	47	60	65	69	72	55
South Carolina	9	18	33	50	54	52
Tennessee	48	56	70	81	91	73

but by county-industry are highly correlated to both the QCEW and CES.

Table **??** shows the state-by-state count of plants with at least twenty employees and a primary SIC code of 3714 from the Dun & Bradstreet. Table 3.10.2 gives the same information, except using data from the County Business Patterns. The two data sets were produced with different methodologies and in different months, but their counts are broadly similar. In a few states and years counts differ by more than a third, but the pattern of plateauing plant counts in northern auto alley and dramatically increasing counts in southern auto alley is seen in both data sets.

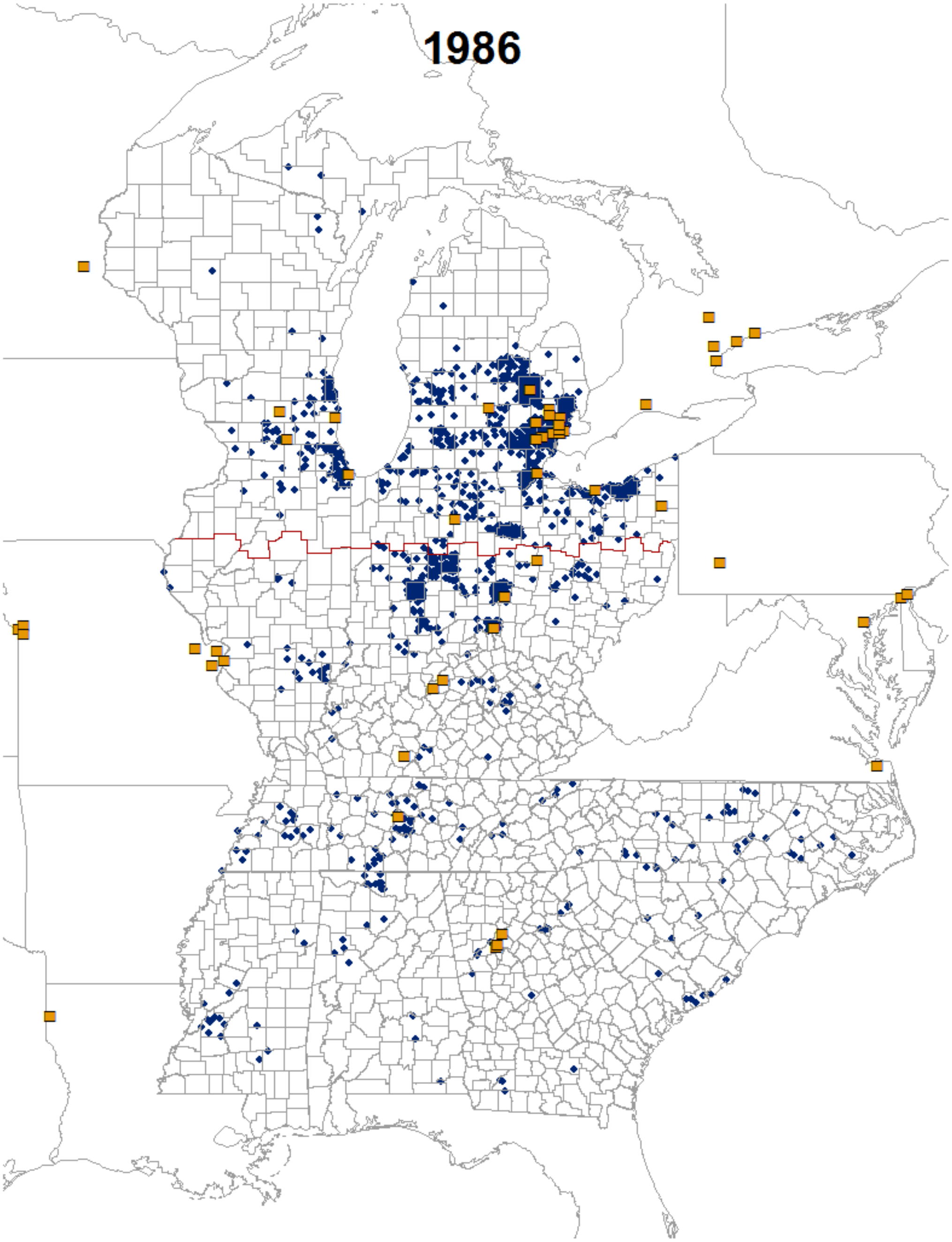


Figure 3.1: Map of supplier employment and assembly plants in 1986.

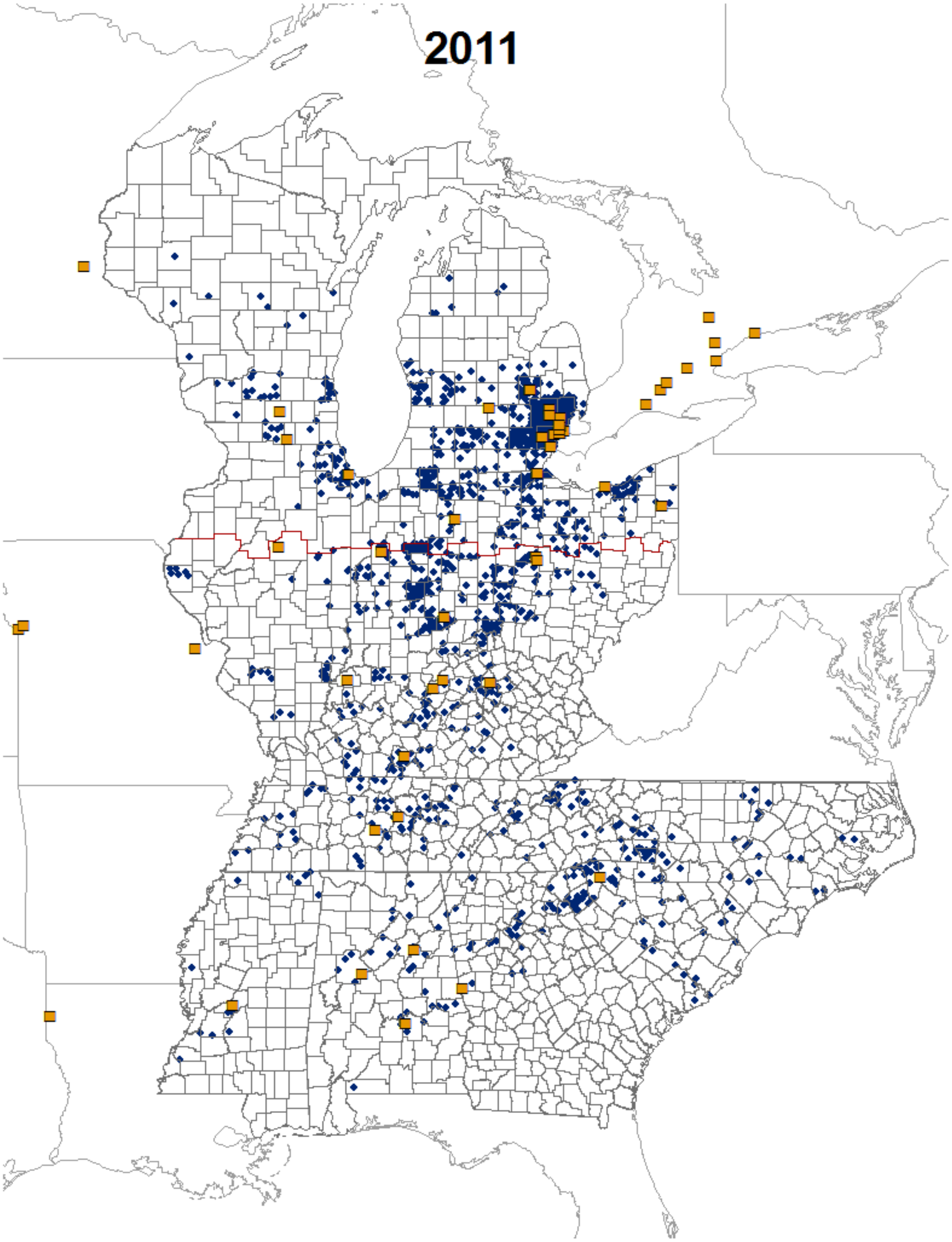


Figure 3.2: Map of supplier employment and assembly plants in 2011.

Chapter 4

Motor Vehicle Assembly Plant Impacts: Evidence from Openings and Closings

4.1 Introduction

Local, state, and provincial governments in North America commonly use subsidies to attract employers to their jurisdictions. Motor vehicle assembly plants are one class of employers that has received particular attention and large subsidies. Policymakers justify the subsidies with claims that assembly plants will spur indirect jobs, particularly by attracting part suppliers. In spite of the enormity of the subsidies in this industry, thorough analysis of what assembly plants actually bring to a community are still needed.

This chapter describes the impact of assembly plant opening. It finds increases in part supplier employment and plant counts are associated with the opening of an assembly plant. Plants are not placed at random locations nor opened at random times. Carmakers generally open plants in places they perceive as especially favorable to manufacturing plants and at a time they perceive as favorable to the industry. Comparisons of the sites assemblers select with the sites they almost select can reveal the casual impact of assembly plant openings. On part supplier employment, this impact is estimated to be positive, but modest: after 8 years an assembly plant brings an average additional 2.0 plants and 408 additional parts supplier jobs to its host county.

This chapter largely adopts the methodology of Greenstone and Moretti (2004) and Greenstone, Hornbeck, and Moretti (2010) using finalists in the site selection process as the control group in a difference-in-differences estimator. This chapter narrows the focus to one industry, motor vehicle manufacturing, so that magnitudes of the impacts found can have clear meaning, and so that the results can inform further analysis of industry, such as chapter 3.

The result here can provide a check on the economic impact models practitioners use (see Loveridge (2004) for a summary and critique of prevailing practices). Connaughton and Madsen (2001) find impact assessments by state development agencies in Alabama and South Carolina that use input-output models and each predict 10,000 jobs resulting from assembly plants they subsequently won. Such predictions contrast with the part supplier location literature (for example, Klier and McMillen (2008) and chapter 2 of this dissertation), which find assembly plants have only a small effect on parts suppliers. The estimates presented in section 4.3 generally fall between the enormous gains predicted by subsidy proponents and the small impact suggested by the locational choice literature.

4.2 Data

Since 1986, 25 new assembly plants have opened in the United States. Some of these plants replaced decommissioned plants. But 15 plants were new investments in an area, opening in a county that had not previously had any car assembly; these new investments are the openings studied in this chapter. With just one exception, all of the openings in new counties (listed in Table 4.1) were by foreign automakers and all of the replacement plants (listed in Table 4.2) were by General Motors, Ford, or Chysler.

The locations of the new plants were carefully. All of the new assembly plants in the United States are in the Midwest or Southeast. Unlike the replacement plants, most of the new openings were in places with relatively small manufacturing sectors and low union membership rates. The process of selecting a location for a new assembly plant is involved. A first stage screens hundreds of potential locations for inclusion on a short list of two to five finalists. The sites on this list sometimes, but not always, become

Table 4.1: US assembly plant openings in new counties, 1986-2011

Assembler	Site	Date of site announcement	Opening Date
General Motors	Roanoke, Indiana	1984	1986
Toyota	Georgetown, Kentucky	1986	1987
Mitsubishi	Normal, Illinois	1985	1988
Subaru	Lafayette, Indiana	1986	1989
Honda	East Liberty, Ohio	1987	1989
BMW	Greer, South Carolina	1992	1994
Toyota	Princeton, Indiana	1995	1996
Daimler	Vance, Alabama	1993	1997
Honda	Lincoln, Alabama	1999	2001
Nissan	Canton, Mississippi	2000	2003
Hyundai	Montgomery, Alabama	2002	2005
Toyota	San Antonio, Texas	2003	2006
Honda	Greensburg, Indiana	2006	2008
Kia Motors	West Point, Georgia	2006	2009
Volkswagen	Chattanooga, Tennessee	2008	2011

public. Assemblers compare specific sites and subsidy offers from the respective state and local governments, then announce the location of their new assembly plant.

For 11 of the new assembly plants, the locations of the other sites that were finalists in assembler's selection process have been made public. Four of the assembly plants are in the sample used in Greenstone and Moretti (2004) and Greenstone, Hornbeck, and Moretti (2010), who find alternate sites printed in trade journal *Site Selection*. Other contemporary news accounts provide lists of finalist sites for 7 more plants. Because multiple finalists are sometimes reported, I have 21 sites that almost were awarded assembly plants. Greenstone and Moretti (2004) and Greenstone, Hornbeck, and Moretti (2010) show that finalist sites have similar observed characteristics as selected sites. This article will join with them in assuming that finalist and selected sites have comparable unobserved characteristics.

Table 4.2: US replacement assembly plant openings, 1986-2011

Assembler	Site	Opening Date
General Motors	Hamtramck, Michigan	1986
AutoAlliance	Flat Rock, Michigan	1987
General Motors	Fairfax, Kansas	1987
General Motors	Lansing, Michigan	1987
Chrysler	Detroit, Michigan	1991
Chrysler	Detroit, Michigan	1995
General Motors	Lansing, Michigan	2001
Chrysler	Toledo, Ohio	2001
Ford	Dearborn, Michigan	2004
General Motors	Lansing, Michigan	2006

Some of the same characteristics that make assembly plants attractive targets for policymakers, their size and durability, limit the number of new openings. A small number of independent observations is therefore inescapable. A broader class of plants would increase offer more openings, but it would be at the expense of the applicability of the impact magnitudes estimated.

The main data set for this chapter is the US Census's County Business Patterns. For each county, County Business Patterns reports the number of plants in each industry and size class. It also reports total employment and total payroll, but these numbers frequently are zeroed out to protect confidentiality when only a few plants of the industry are in a county. I construct an estimate of the censured employment number by multiplying the number of plants in each size class with the average employment of that size class for that industry in that state (or nation, if state employment for the size class is also censured), but this introduces noise, and so I also use counts of plants with 20 or more employees. County Business Patterns classified industries according to four digit Standard Industry Classification (SIC) codes through 1998 and switched to six digit North American Industry Classification System (NAICS) codes beginning in 1999. When reporting industry specific results, I use SIC codes and sum over corresponding

Table 4.3: Growth in new counties hosting assembly plants in the 8 years following the site selection announcement

Variable	Obs	Mean	Std. Dev.	Min	Max
All industries					
Change in employment	10	15,333	20,371	-3,590	69,848
Change in plant count	10	503.3	761.6	-298	2,339
Change in plants with 20+ employees	10	132.0	204.6	-58	649
Parts suppliers (SIC 3714)					
Change in employment	10	203	328	0	1,032
Change in plant count	10	1.9	2.3	0	6
Change in plants with 20+ employees	10	1.1	1.6	0	5

NAICS codes in later years.

4.3 Evidence from new openings

Simple averages in the raw data hint that new assembly plants increase employment in their host counties. Total employment grows on average by 15,333 (on an average base of 94,322) in the eight years following the site selection announcement. The number of part supplier plants (SIC 3714) increases by 1.7, the number of part supplier plants with 20 or more employees increases by 1.0, and estimated part supplier employment increases by an average of 202. Table 4.3 shows that the averages are coupled with high variance; some of the new plant sites saw little movement as the assembly plant opened. (Plants announced after 2003 are not included, as 8 years had not elapsed by the last year of released data.) Table 4.4 displays the evolution of part supplier plant counts in each of the new greenfield plants.

The growth in supplier employment is more impressive given the contraction in supplier employment (and manufacturing employment generally) during this time frame. Yet carmakers open assembly plants more when they expect demand for cars and the size of the industry to grow. Assembly plants are not placed at random, but rather in places attractive to manufacturers, so growth rates or simple comparisons with national

Table 4.4: Counts of supplier plants (SIC 3714) with 20+ employees in counties before and after new assembly plants open

		Plants within county				
Assembler	County	6 yrs prior	3 yrs prior	at opening	3 yrs after	6 years after
Hyundai	Montgomery Co., AL	1	2	4	10	6
Honda	Talladega Co., AL	0	1	1	1	2
Daimler	Tuscaloosa Co., AL	0	0	1	0	1
Kia	Troup Co., GA	2	2	5		
Mitsubishi	McLean Co., IL			0	1	2
GM	Allen Co., IN			7	9	12
Honda	Decatur Co., IN	3	3	1	3	
Toyota	Gibson Co., IN	0	0	0	0	0
Subaru	Tippecanoe Co., IN		1	0	0	3
Toyota	Scott Co., KY			0	0	1
Toyota	Madison Co., MS			3	1	1
Honda	Logan Co., OH		1	1	2	3
BMW	Spartanburg Co., SC	6	6	9	9	7
Volkswagen	Hamilton Co., TN	2	4	2		
Toyota	Bexar Co., TX	10	7	9	13	

averages do not give the structural impact of attracting an assembly plant.

The sites that were finalist in assembler's selection process can be used as a control group for comparison. In the eight years following the announcement that they would not get an assembly plant, counties that were runners-up in the selection process added an average 65 motor vehicle parts supplier jobs, but lost on average 0.3 supplier plants. An estimate of the causal effect of a new assembly plant on parts supplier employment 8 years after the announcement would be 203-65=138 jobs. If there was exactly one runner-up site for every assembly plant, then this simple difference in differences would be sufficient.

4.3.1 Econometric Model and Results

This section presents a simple regression model that is closely related to this difference-in-differences approach, following Greenstone and Moretti (2004). Let $y_{\ell t}$ be the outcome variable - employment or plant count - in location ℓ at year t. Let X_ℓ be 1 if the location is selected for an assembly plant and 0 if it is a finalist that does not get an assembly plant. Let the announcement date for the location be denoted d_ℓ. Assume

$$y_{\ell t} = \sum_{\tau=-25}^{25} \beta_{S\tau} I[X_\ell = 1] I[t - d_j = \tau] + \sum_{\tau=-25}^{25} \beta_{N\tau} I[X_\ell = 0] I[t - d_j = \tau] I[t - d_\ell = \tau] + \gamma_\ell + \delta_t + \epsilon_{\ell t}$$

where $I[\cdot]$ is the indicator function, γ_ℓ is a county-specific fixed effect, δ_t is a common time effect, and $\epsilon_{\ell t}$ is an independent, identically distributed error term. The coefficient $\beta_{S\tau}$ will give the average outcome conditional on the location being selected for an assembly plant announced τ years previously; $\beta_{S\tau}$ will give the average outcome conditional on the location being a finalist not selected for an assembly plant announced τ years previously. The difference will identify the average causal impact of adding an assembly plant for each year after the announcement. An equivalent regression equation

$$y_{\ell t} = \sum_{\tau=-25}^{25} \beta_{S\tau} I[t - d_j = \tau] - \sum_{\tau=-25}^{25} (\beta_S \tau - \beta_{N\tau}) I[X_\ell = 0] I[t - d_\ell = \tau] + \gamma_\ell + \delta_t + \epsilon_{\ell t}$$

is estimated to find the standard errors on the differences, $\beta_S \tau - \beta_{N\tau}$.

To omit the time location and time fixed effects (γ_ℓ and δ_t) is to calculate simple means conditional on the location being selected or not. The assembly plant announcements have varying numbers of alternative sites, so the announcement dates for selected locations and rejected locations have a different temporal distribution. The year fixed effect prevents bias that might arise if an assembly plant with a longer list of alternatives was announced at an extreme of the business cycle. Another approach would be to randomly select one runner-up from the finalist list for each assembly plant, but to do so would reduce the already small number of observations. The location fixed effect γ_ℓ is necessary whenever the panel is not balanced, because counties have different sizes and initial values.

Table 4.5 displays results for two outcome variables: parts supplier employment in the host county and the number of parts supplier plants with at least 20 employees. I normalize β_{S0} and β_{N0} to 0 so that in subsequent years the coefficients measure the conditional average gains since the assembly plant site announcement. Taken alone, the estimates for $\beta_{S\tau}$ in columns 1 and 4 would be a naive estimator that accounts for time-effects and the unbalanced panel, but ignores propensity of assembly plants to locate in favorable sites. The structural estimator $\beta_{S\tau} - \beta_{N\tau}$, in columns 3 and 6, is lower by a third to half for most years following the plant announcement.

Note that $\beta_{S\tau} - \beta_{N\tau}$ are often not significant. This is not surprising given the paucity of observations and variability within the data. The conclusions that winning assembly plant increase supplier employment or brings in new supplier plants cannot be made with confidence using a generation's worth of data. The results, however, do have enough statistical power to reject that the average effect of an assembly plant is to bring several thousand indirect parts supplier jobs to a county. The estimates for the causal impact of an assembly plant are more modest: 408 more parts supplier jobs after 8 years in 1.97 additional plants.

Table 4.6 shows results using total employment in the host county as the outcome variable. The employment at the assembly plant itself is included as well as indirect jobs created at parts suppliers or any other sector. From four to eight years after the announcement, the assembly plant seems to cause an increase in employment of around 10,000. These results, however, are not statistically significant and are not robust to dropping single assembly plants. The drop seen between eight and nine years after the announcement is primarily a reflection of data for the Toyota plant in San Antonio being present for year eight but not year nine.

Table 4.5: Effect of Assembly Plant Selection on Parts Suppliers in the same county

	Plants with 20 or more employees			Industry employment		
	$\beta_{S\tau}$	$\beta_{N\tau}$	$\beta_{S\tau} - \beta_{N\tau}$	$\beta_{S\tau}$	$\beta_{N\tau}$	$\beta_{S\tau} - \beta_{N\tau}$
$\tau = -10$	-2.958***	-2.578***	-0.379	-719.1*	-923.3***	204.1
	(0.812)	(0.616)	(0.753)	(306.4)	(232.6)	(284.4)
$\tau = -5$	-1.824**	-1.162*	-0.662	-211.4	-522.6**	311.2
	(0.611)	(0.475)	(0.663)	(230.6)	(179.2)	(250.1)
$\tau = -4$	-1.616**	-0.657	-0.959	-115.6	-363.0*	247.4
	(0.606)	(0.456)	(0.681)	(228.9)	(172.0)	(257.2)
$\tau = -3$	-0.998	-0.242	-0.756	-130	-322.4	192.3
	(0.565)	(0.439)	(0.663)	(213.4)	(165.9)	(250.3)
$\tau = -2$	-0.832	-0.0524	-0.779	-138.5	-293.3	154.9
	(0.552)	(0.433)	(0.662)	(208.2)	(163.5)	(249.8)
$\tau = -1$	-0.486	0.0771	-0.563	-80.14	-7.248	-72.89
	(0.521)	(0.423)	(0.652)	(196.6)	(159.8)	(246.1)
$\tau = 0$	0	0	0	0	0	0
$\tau = 1$	-0.124	0.185	-0.309	-22.71	-165.1	142.4
	(0.502)	(0.388)	(0.630)	(189.4)	(146.5)	(237.7)
$\tau = 2$	0.0341	0.375	-0.341	-51.18	-83.16	31.98
	(0.482)	(0.405)	(0.626)	(181.7)	(152.8)	(236.1)
$\tau = 3$	0.726	0.68	0.0455	240.1	25.88	214.2
	(0.464)	(0.432)	(0.631)	(175.0)	(163.0)	(238.3)
$\tau = 4$	1.776***	0.930*	0.846	527.9**	-41.93	569.8*
	(0.462)	(0.462)	(0.636)	(174.5)	(174.5)	(240.2)
$\tau = 5$	1.860***	1.103*	0.757	185.3	127.8	57.52
	(0.473)	(0.477)	(0.641)	(178.7)	(180.1)	(242.0)
$\tau = 6$	2.494***	1.156*	1.338*	764.8***	372.7	392.2
	(0.498)	(0.512)	(0.667)	(188.0)	(193.2)	(251.8)
$\tau = 7$	3.238***	1.197*	2.041**	483.6*	349.7	133.9
	(0.509)	(0.530)	(0.661)	(192.2)	(200.0)	(249.6)
$\tau = 8$	3.186***	1.213*	1.973**	640.7**	232.6	408.0
	(0.522)	(0.535)	(0.653)	(196.9)	(202.0)	(246.6)
$\tau = 9$	4.149***	1.672**	2.477***	774.6***	503.3*	271.3
	(0.553)	(0.608)	(0.707)	(208.8)	(229.4)	(267.0)
$\tau = 10$	3.655***	1.819**	1.836*	905.0***	369.9	535.1
	(0.593)	(0.681)	(0.760)	(223.8)	(256.9)	(286.8)
$\tau = 15$	5.456***	3.363***	2.093**	1024.4***	673.6*	350.9
	(0.716)	(0.806)	(0.740)	(270.1)	(304.3)	(279.4)
$\tau = 20$	7.266***	4.610***	2.656**	1592.3***	1376.1***	216.2
	(0.926)	(1.068)	(0.923)	(349.6)	(403.0)	(348.3)

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.6: Effect of Assembly Plant Selection on Total Employment in the Same County

	$\beta_{S\tau}$	$\beta_{N\tau}$	$\beta_{S\tau} - \beta_{N\tau}$
$\tau = -10$	-21454.9	1850.4	-23305.3*
	(12273.6)	(10688.1)	(11470.1)
$\tau = -5$	-7838.0	-3005.0	-4833.0
	(9246.0)	(8118.5)	(10413.6)
$\tau = -4$	-4930.0	-2509.3	-2420.7
	(8851.7)	(7779.3)	(10431.6)
$\tau = -3$	-1441.0	-2449.0	1008.0
	(8498.0)	(7416.8)	(10409.1)
$\tau = -2$	1042.3	-1489.2	2531.5
	(8243.6)	(7367.0)	(10424.7)
$\tau = -1$	-305.3	-360.7	55.42
	(7810.7)	(7223.9)	(10277.4)
$\tau = 0$	0	0	0
$\tau = 1$	-227.0	203.1	-430.2
	(7212.0)	(6806.3)	(9789.8)
$\tau = 2$	-450.2	442.6	-892.8
	(7036.0)	(6952.5)	(9717.9)
$\tau = 3$	3647.4	910.1	2737.3
	(7049.9)	(7016.9)	(9685.5)
$\tau = 4$	7899.4	-326.0	8225.4
	(7325.3)	(7482.3)	(9897.5)
$\tau = 5$	9550.3	-1547.1	11097.4
	(7452.9)	(7715.1)	(9883.5)
$\tau = 6$	13787.4	807.6	12979.8
	(8012.1)	(8694.8)	(10606.1)
$\tau = 7$	12744.4	1236.2	11508.2
	(8290.7)	(9051.2)	(10587.4)
$\tau = 8$	13776.5	1160.4	12616.1
	(8587.0)	(9333.5)	(10587.2)
$\tau = 9$	4519.2	3930.9	588.3
	(9147.9)	(10778.2)	(11574.4)
$\tau = 10$	7410.0	5336.6	2073.4
	(9815.8)	(11611.9)	(12068.8)
$\tau = 15$	9903.0	4798.9	5104.1
	(12710.7)	(14237.0)	(12549.2)
$\tau = 20$	9326.0	5878.5	3447.5
	(16674.3)	(18823.1)	(15959.2)

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3.2 Distance Range of Assembly Plant Effects

How far spillovers from an assembly plant extend matters to policymakers. If assembly plants promote employment only in their immediate neighborhoods, then state subsidies serve to redistribute wealth between regions within a state. If assembly plants promote employment in a much wider radius, then the state subsidizing the assembly plants is not capturing all of the benefits, some of which accrue to neighboring states.

The same difference-in-differences methodology is used where the outcome variables are the number of large part suppliers plants and employment in circles of various radii around the assembly plants and its runner-up sites. Table 4.7 shows the results for large plant counts and table 4.8 for employment. Doubling the radius of the study regions quadruples their area. The effect on plant counts increases with distance, but at a slower, diminishing rate. Nearly all of the additional part supplier plants an assembly plant attracts will locate within 200 kilometers. For a medium sized state like Alabama, a plant opening near the center of the state could be expected to bring parts supplier plants throughout the state, yet the state would internalize most of the benefits from its subsidy.

4.4 Conclusion

With so few assembly plant openings, the effects of new assembly plants on their surroundings cannot be identified with precision. There is enough statistical power to assert that assembly plants cause some additional suppliers to nearby. The effect can be overstated by analysis that ignores how the growth that would come even without an assembly plant at locations so favorable to manufacturing. It can be understated by analysis that ignores a background of decreasing employment industry-wide. From the point estimates, the impact of assembly plants still appears substantial: dozens of additional supplier plants and thousand of jobs within 200 kilometers.

Table 4.7: Effect of Assembly Plant Selection on Parts Suppliers with at least 20 employees

$(\beta_{S\tau} - \beta_{N\tau})$	in same county	within 100 km	within 200 km	within 400 km
$\tau = -10$	-0.379	-0.913	-3.748	3.770
	(0.753)	(3.275)	(7.794)	(16.92)
$\tau = -5$	-0.662	0.977	-1.35	7.446
	(0.663)	(2.946)	(7.012)	(15.22)
$\tau = -4$	-0.959	2.138	-0.0319	6.017
	(0.681)	(2.946)	(7.011)	(15.22)
$\tau = -3$	-0.756	1.157	0.546	7.407
	(0.663)	(2.934)	(6.984)	(15.16)
$\tau = -2$	-0.779	1.39	2.106	11.09
	(0.662)	(2.946)	(7.011)	(15.22)
$\tau = -1$	-0.563	0.798	0.697	12.42
	(0.652)	(2.944)	(7.007)	(15.21)
$\tau = 0$	0	0	0	0
$\tau = 1$	-0.309	1.533	6.25	7.508
	(0.630)	(2.752)	(6.551)	(14.22)
$\tau = 2$	-0.341	0.123	5.335	5.501
	(0.626)	(2.747)	(6.491)	(14.09)
$\tau = 3$	0.0455	1.929	8.27	10.50
	(0.631)	(2.738)	(6.471)	(14.04)
$\tau = 4$	0.846	4.484	12.55	16.31
	(0.636)	(2.803)	(6.619)	(14.37)
$\tau = 5$	0.757	4.702	13.66^{*}	15.62
	(0.641)	(2.804)	(6.620)	(14.37)
$\tau = 6$	1.338^{*}	5.235	13.03	15.92
	(0.667)	(2.980)	(7.092)	(15.39)
$\tau = 7$	2.041^{**}	6.852^{*}	15.69^{*}	20.97
	(0.661)	(2.978)	(7.087)	(15.38)
$\tau = 8$	1.973^{**}	6.057^{*}	16.36^{*}	17.13
	(0.653)	(2.983)	(7.101)	(15.41)
$\tau = 9$	2.477^{***}	5.591	17.23^{*}	28.80
	(0.707)	(3.248)	(7.730)	(16.78)
$\tau = 10$	1.836^{*}	7.351^{*}	18.44^{*}	26.89
	(0.760)	(3.395)	(8.079)	(17.53)
$\tau = 15$	2.093^{**}	7.098^{*}	24.22^{**}	28.48
	(0.740)	(3.472)	(8.262)	(17.93)
$\tau = 20$	2.656^{**}	7.196	18.38	47.92^{*}
	(0.923)	(4.397)	(10.46)	(22.71)

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.8: Effect of Assembly Plant Selection on Parts Supplier Employment

$(\beta_{S\tau} - \beta_{N\tau})$	in same county	within 100 km	within 200 km	within 400 km
$\tau = -10$	204.1	2314.2	-1041.4	-7010.1
	(284.4)	(2086.5)	(5728.7)	(11306.2)
$\tau = -5$	311.2	986	380.9	-7705.7
	(250.1)	(1877.1)	(5154.0)	(10172.0)
$\tau = -4$	247.4	276	-585.9	-8550.5
	(257.2)	(1876.9)	(5153.5)	(10171.0)
$\tau = -3$	192.3	915.4	978.7	-3251.3
	(250.3)	(1869.6)	(5133.4)	(10131.4)
$\tau = -2$	154.9	691.8	-84.61	-4509.1
	(249.8)	(1876.8)	(5153.1)	(10170.2)
$\tau = -1$	-72.89	86.32	504.1	-1760.4
	(246.1)	(1875.9)	(5150.7)	(10165.4)
$\tau = 0$	0	0	0	0
$\tau = 1$	142.4	1385.4	1672.4	-9572.5
	(237.7)	(1753.5)	(4814.8)	(9502.5)
$\tau = 2$	31.98	723.7	2842.1	-668.5
	(236.1)	(1750.0)	(4771.1)	(9416.2)
$\tau = 3$	214.2	427.7	-689.4	-9609.3
	(238.3)	(1744.6)	(4756.2)	(9387.0)
$\tau = 4$	569.8*	2332.5	4453.0	-6865.7
	(240.2)	(1786.2)	(4865.1)	(9601.8)
$\tau = 5$	57.52	-449.3	-1634.5	-8678.3
	(242.0)	(1786.5)	(4865.5)	(9602.7)
$\tau = 6$	392.2	1446.7	1290.7	-5024.0
	(251.8)	(1898.5)	(5212.4)	(10287.2)
$\tau = 7$	133.9	573.1	2385.7	-4731.9
	(249.6)	(1897.3)	(5209.0)	(10280.6)
$\tau = 8$	408.0	1244.5	5377.3	3212.6
	(246.6)	(1900.9)	(5219.0)	(10300.3)
$\tau = 9$	271.3	1366.7	7044.1	5543.9
	(267.0)	(2069.6)	(5681.7)	(11213.5)
$\tau = 10$	535.1	2518.1	7967.5	7078.5
	(286.8)	(2163.1)	(5937.9)	(11719.1)
$\tau = 15$	350.9	-464.2	1591.6	-2108.3
	(279.4)	(2212.2)	(6072.6)	(11984.9)
$\tau = 20$	216.2	-1457.8	-4124.5	-12926.9
	(348.3)	(2801.4)	(7690.6)	(15178.2)

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Bibliography

Aguirregabiria, V., and P. Mira (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75(1), 1–53.

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

Benkard, L., A. Bodoh-Creed, and J. Lazarev (2010): "Simulating the Dynamic Effects of Horizontal Mergers: U.S. Airlines," Working paper, Yale University.

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation.," *RAND Journal of Economics*, 25(2), 242 – 262.

Brinkman, J., D. Coen-Pirani, and H. Sieg (2012): "Estimating a Dynamic Equilibrium Model of Firm Location Choices in an Urban Economy," Working paper, Federal Reserve Bank of Philadelphia.

Caplin, A., and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), pp. 25–59.

Carlton, D. W. (1979): "Why New Firms Locate Where They Do: An Econometric Model," in *Interregional Movements and Regional Growth*, ed. by W. C. Wheaton, pp. 13–50. Urban Institute, Washington.

Chintagunta, P., J.-P. Dubé, and V. Singh (2003): "Balancing Profitability and Customer Welfare in a Supermarket Chain," *Quantitative Marketing and Economics*, 1(1), 111–147.

COLLARD-WEXLER, A. (2010): “Mergers and Sunk Costs: Am application to the ready-mix concrete industry,” Working paper, NYU Stern.

CONNAUGHTON, J. E., AND R. A. MADSEN (2001): “Assessment of Economic Impact Studies: The Cases of BMW and Mercedes-Benz,” *The Review of Regional Studies*, 31(3).

CORTS, K. S. (1998): “Third-degree price discrimination in oligopoly: all-out competition and strategic commitment.,” *RAND Journal of Economics*, 29(2), 306 – 323.

DUMAIS, G., G. ELLISON, AND E. L. GLAESER (2002): “Geographic Concentration as a Dynamic Process,” *Review of Economics and Statistics*, 84(2), 193–204.

ELLISON, G., AND E. L. GLAESER (1997): “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach,” *Journal of Political Economy*, 105(5), 889–927.

ELLISON, G., E. L. GLAESER, AND W. R. KERR (2010): “What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns,” *American Economic Review*, 100(3), 1195–1213.

ERICSON, R., AND A. PAKES (1995): “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, 62(1), 53–82.

GANDHI, A., Z. LU, AND X. SHI (2013): “Estimating demand for differentiated products with error in market shares,” Discussion paper, University of Wisconsin - Madison.

GREENSTONE, M., R. HORNBECK, AND E. MORETTI (2010): “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings,” *Journal of Political Economy*, 118(3), 536–598.

GREENSTONE, M., AND E. MORETTI (2004): “Bidding for Industrial Plants: Does Winning a ‘Million Dollar Plant’ Increase Welfare?,” Working paper, Massachusetts Institute of Technology.

Head, K., J. Ries, and D. Swenson (1995): “Agglomeration benefits and location choice: Evidence from Japanese maunfacturing investments in the United States,” *Journal of International Economics*, 38(3-4), 223–247.

Hirsch, B. T., and D. A. MacPherson (2003): “Union Membership and Coverage Database from the Current Population Survey: Note,” *Industrial and Labor Relations Review*, 56(2), 349–354.

Holmes, T. J. (1989): “The Effects of Third-Degree Price Discrimination in Oligopoly.,” *American Economic Review*, 79(1), 244.

——— (1999): “How Industries Migrate When Agglomeration Economies Are Important,” *Journal of Urban Economics*, 45(2), 240–263.

——— (2004): “Step-by-step migrations,” *Review of Economic Dynamics*, 7(1), 52–68.

Klier, T., and D. P. McMillen (2008): “Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples,” *Journal of Business & Economic Statistics*, 26(4), 460–471.

Loveridge, S. (2004): “A Typology and Assessment of Multi-sector Regional Economic Impact Models,” *Regional Studies*, 38(3), 305–317.

Miller, N. J., and M. Osborune (2011): “Competition Among Spatially Differentiated Firms: An Estimator with an Application to Cement,” *Working Paper*.

Montgomery, A. L. (1997): “Creating micro-marketing pricing strategies using supermarket scanner data.,” *Marketing Science*, 16(4), 315.

Neumark, D., B. Wall, and J. Zhang (2011): “Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series,” *Review of Economics and Statistics*, 93(1), 16–29.

Nevo, A. (2001): “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 69(2), 307–342.

Pakes, A., M. Ostrovsky, and S. Berry (2007): “Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry-Exit Examples),” *RAND Journal of Economics*, 38(2), 373–399.

PETRIN, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4).

ROSENBAUM, T. (2012): "Where Do Automotive Suppliers Locate and Why?," Working paper, Yale University.

RUBENSTEIN, J. M. (1992): *The Changing US Auto Industry: a geographical analysis.* Routledge, London.

RYAN, S. P. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 80(3), 1019–1061.

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

SMITH, D., AND R. FLORIDA (1994): "Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturers in Automotive-related Industries," *Journal of Urban Economics*, 36(1), 23–41.

SU, C.-L., AND K. L. JUDD (2012): "Constrained optimization approaches to estimation of structural models," *Econometrica*, 80(5), 2213–2230.

VAN BIESEBROECK, J. (2003): "Productivity Dynamics with Technology Choice: An Application to Automobile Assembly," *Review of Economic Studies*, 70(1), 167–198.

WOODWARD, D. (1992): "Locational Determinants of Japanese Manufacturing Startups in the United States," *Southern Economic Journal*, 58(3), 690–708.

ZBARACKI, M. J., M. RITSON, D. LEVY, S. DUTTA, AND M. BERGEN (2004): "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *The Review of Economics and Statistics*, 86(2), 514–533.