Migration and Agglomeration Among Motor Vehicle Parts Suppliers

Brian Adams

March 3, 2012

Abstract

Motor vehicle manufacturing in North America has been geographically clustered for most of its history, but the agglomeration forces that maintained such clustering have not prevented an industry migration in recent decades. As production migrates, states have offered large subsidies to new assembly plants in hopes of also attracting a cluster of suppliers. This paper uses supplier and assembly plant locations from 1986 to 2011 to estimate a dynamic structural model of suppliers' site selection and plant closure decisions. The model contains terms for various local costs and for agglomeration effects. It models future expectations of costs and assembler locations, which otherwise would bias static models. Finally, this paper estimates the locations of suppliers under counter-factual placement of assembly plants; this provides case-specific estimates for the additional employment gains a successful assembly plant subsidy bid generates.

1 Introduction

Production in many industries is geographically concentrated. Some industries, including motor vehicle manufacturing, are far more clustered than if plants were distributed according to population or placed at random. Theory long has provided a list of potential agglomeration mechanisms, but empirical work is only beginning to examine the relative importance of the various agglomeration mechanisms. This paper uses the plant location decisions of automobile parts suppliers to determine what is responsible for clustering in motor vehicle parts manufacturing. It thereby also informs local government policies of subsidizing assembly plants in order to attract suppliers.

North American motor vehicle manufacturing, especially motor vehicle parts manufacturing, has been concentrated around the lower Great Lakes for most of its history. The agglomeration forces that maintained such clustering have not prevented dramatic industry migrations in recent decades. Final assemblers that had plants near many large markets have retreated back to the Midwest. Other assemblers established plants in the South. Parts suppliers have spread southward.

The migration, of course, is composed of individual firm decisions on where to locate a new plant or whether to close an existing plant. This paper builds a structural model of these site selection and plant closure decisions. The location of a plant determines what customers are near, what transportation costs may be, from what labor markets the firm may hire, and from what productivity spillovers it may benefit. Entrants in the model to be presented chose where to locate and when to close based on these variables. Because plants are durable, expectations of future location characteristics and assembly plant placement affect supplier's decision. This model of entry and exit therefore incorporates dynamics and future expectations.

This model generalizes the traditional approach to modeling the site selection of parts suppliers. Setting the discount factor in the model here to zero and only using entry decisions in the dynamic model will yield a model closely related to the existing literature for this industry. The full model brings in exit decision data and corrects a bias the migration introduces into static models. Forward-looking suppliers respond not only to a location's current characteristics, but also to their expectation of its future characteristics. A firm sensitive to transportation costs may locate far away from existing assembly plants if it believes assemblers will soon migrate to its current location. If that entry decision is modeled with only contemporaneous variables, the importance of distance would be underestimated.

The full model joins a growing literature that brings spatial considerations into the latest empirical industrial organization models. Much as Thomadsen (2005) brings transportation costs into demand estimation and Seim (2006) brings transportation costs into a static entry model, this paper brings geography and transportation costs into a dynamic entry model. The primary challenge in doing this is the difficulty of maintaining state variables for each location. The paper first reduces the number of state variables by summarizing distances and associated transportation costs with one scalar index for each location. Then, by assuming new entrants but not incumbents are mobile, the estimation approach breaks the problem into separate single-agent decision problems for each county.

The industry migration has lead state and local governments to compete for new assembly plants. State and local governments routinely offer carmakers tax breaks, infrastructure improvements, job training, and other incentives worth hundreds of millions of dollars to select them as a location for a new assembly plant. Justifications for such incentives often appeal to the ability of a new assembly plant to attract suppliers (Rubenstein 2008). The dynamic model is capable of experiments estimating the placement of suppliers in the counterfactual cases where assembly plants go to an alternative location instead of the site actually selected. Such counterfactuals give case-specific estimates of what successful subsidy bids brought.

Many of the common theorized causes of industry clustering appear in the model as location specific variables in the suppliers' cost functions. Firms may locate in the same area to be near the same customers. Tier one auto suppliers have a specific set of customers, assemblers, with known locations and order quantities. Thus the distance to customers and potential customers is known with precision. Firms might locate in clusters because productivity or cost advantages come from proximity to similar firms. These advantages could come from technological spillovers or from having a larger labor pool. The number of other suppliers nearby is known. Finally, firms might locate in the same area not to be near other firms, but because they all are attracted by the same "natural advantages" of the area. For auto suppliers, these may be observable variables such as local manufacturing wages, attitudes towards unionization, or population, all of which are included in the model. The relative importance of these clustering mechanisms will be measured by the estimated model parameters.

This work complements other recent work on the causes of clustering. Greenstone, Hornbeck, and Moretti (2010) provide evidence for the existence of local productivity spillovers and Ellison, Glaeser, and Kerr (2010) use the coagglomeration of related industries to show that proximity to customers, proximity to suppliers, technological spillovers, labor pooling, and natural advantages all have some roles in agglomeration. Which forces are most important may differ from industry to industry. The structural model in this paper determines the relative power of several agglomeration factors for a specific industry. The counterfactual predictions performed underscore the importance of understanding which variables are most responsible for clustering.

2 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 usually included in broader definitions of auto alley, namely: Wisconsin, Michigan, Illinois, Indiana, Ohio, Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Alabama, and Mississippi. In 2009, these states accounted for 75% of national employment in motor vehicle parts manufacturing. In section 3 and in some tables, I refer to Michigan, Wisconsin, and the portions of Illinois, Indiana, and Ohio north of 42.5° latitude as "northern auto alley" and Kentucky, Tennessee, North Carolina, South Carolina, Georgia, Alabama, Mississippi, and the rest of Illinois, Indiana, and Ohio as "southern auto alley."

Supplier plant locations are based on Dun & Bradstreet data. Plants are matched through time to create a panel.[1] Entry is defined as the presence of a plant not operating in any previous period. Closure is defined as a plant not present in the any future period. Estimates use plants with a primary Standard Industrial Classification code of 3714. Early Dun & Bradstreet data have been criticized for inflated employment counts and delayed recognition of entry. These problems do not seem pronounced for the industry and time period used here. Appendix A2 discusses this further.

For estimation, a county is a location. Location characteristics include distances to every assembly plant, local manufacturing unionization rates, interstate highway presence, and an evolving manufacturing wage. Appendix A1 lists specific sources for each variable.

[1] Details of linking the cross sections are described in Appendix A1.

3 Industry Migration

For much of the North American automobile industry's history, suppliers were highly concentrated in Michigan and the lower Great Lakes region. Final assembly was more market-oriented with Ford and General Motors (GM) operating branch assembly plants near major markets throughout North America. With declining market share and an increasing number of models, economies of scale prevented assembly plants from making a large share of the product line near final costumers. Beginning in the 1980s, the final assembly for the American brands retreated to the Midwest. In 1986, Ford and GM operated 24 assembly plants in the United States outside of northern auto alley; in 2011 that number was 12.

Contemporaneously, automakers headquartered in Asia and Europe built "transplant" assembly plants in the Midwest and South. Volkswagen opened an assembly plant in western Pennsylvania in 1978, Honda followed in central Ohio in 1982, Nissan entered in Tennessee in 1984. Toyota began joint operations 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 States, [2] of which 13 are in the southern auto alley. Because of transplant construction (and despite Ford and GM closures in Atlanta), the number of assembly plants in the southern auto alley has more than doubled since the early 1980s.

Suppliers have also migrated. From 1986 to 2011, in the southern auto alley both motor vehicle parts employment and number of plants doubled. Meanwhile, in the northern auto alley the number of plants increased by only 38%; parts employment had almost returned to its 1986 level by 2011 after peaking in 2001. The sites selected by new entrants on average were further south and had lower union membership than the sites of incumbent suppliers, as detailed in table 1.

Supplier plant turnover has been considerable in all regions. Even during the South's dramatic growth, a third of all incumbent plants would close in a five-year period. Despite the general emigration of parts suppliers from the North since 2001, hundreds of new plants have entered the northern auto alley. Table 2 describes the evolution of the number of plants by state and region; table 3 the evolution of new plants.

4 Model

This paper uses a model of supplier plant entry and exit. Supplier plants and potential entrants are players in a dynamic game like those studied in Ericson and Pakes (1995). Final assembly plants are the source of demand for parts suppliers, but assembly plant placement and production quantities are taken as exogenous.[3]

[2] The number includes Mazda's joint venture with Ford. Ontario also hosts 1 joint venture and 3 transplant assembly plants.

[3] For a theoretical model in which assemblers and suppliers location decisions are interdependent, see Holmes (2004). Exogenously placed assembly plants will be help in the counterfactual, where the effects of alternate assembly plant placements are explored.

Time is discrete in this model. 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 all other suppliers. After a new entrant selects its location and becomes an incumbent supplier, its only remaining decision is when to exit.

Both the revenues and costs for suppliers depend on location. Let locations for supplier plants be indexed by $\ell = 1, ..., L$, let the number of suppliers at ℓ be $n_{\ell t}$, and let $X_{\ell t}$ be a vector of location-specific characteristics. Index assembly plants by $a = 1, ..., A$ and let their output quantities in period t be q_{at}. Denote the distance between assembly plant a and location ℓ by $d_{\ell a}$. Revenues for a supplier plant may depend on the location characteristics and proximity to each assembly plant of all competing supplier plants. Define a vector of supplier counts $n_t = \{n_{1t}, \ldots, n_{Lt}\}$ and a matrix of location characteristics $X_t = \{X_{1t}, \ldots, X_{Lt}\}$.

Suppliers earn revenue by selling parts to assembly plants. The function $s_{\ell a}(n_t, X_t)$ gives the share of parts used at assembly plant a from any given plant in location l as its supplier for a part. (Alternatively, $s_{\ell a}$ could be interpreted as the probability of a plant in ℓ being selected as the sole supplier for assembly plant a.) Since suppliers compete, the probability of supplying each assembly plant is a function of all supplier locations and local characteristics. Each assembly plant selects a supplier by a process that can be modeled as a multinomial logit, so

$$s_{\ell a}(n_t, X_t) = \frac{e^{\delta X_{\ell t} + \delta_{distance} d_{\ell a}}}{\sum_j n_{jt} e^{\delta X_{jt} + \delta_{distance} d_{ja}}}.$$

The revenue for supplying assembly plant a is $\gamma_p q_{at} s_{\ell a}(n_t, X_t)$, so the total revenue for a supplier at location ℓ is $\sum_{a=1}^{A} \gamma_p q_{at} s_{\ell a}(n_t, X_t)$.

Production costs at a location are a function of local characteristics, $C(X_{\ell t}, n_{\ell t}) = -\gamma_X X_{\ell t} - \gamma_n n_{\ell t}$. To allow for local productivity spillovers, costs depend on the number of other suppliers in the location. Local production costs also depend on location characteristics X that include population density, interstate highway presence, local manufacturing wages, unionization rates, and period indicators. Distances to assembler plants do not contribute to production costs in this model, since industry practice is for the assembler or for logistics firms hired by the assembler to arraign most shipping. Hence the supplier cost function C is the same function for every location and does not directly depend on variables from any other location. (In contrast, the market share functions $s_{\ell a}$ are indexed by ℓ, because distances are different for every location and only the ℓth component of X and n appear in the numerator.) 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).)

Supplier plants in location ℓ earns period profits

$$\pi_\ell(n_t, X_t, q_t) = \sum_{a=1}^{A} \gamma_p q_{at} s_{\ell a}(n_t, X_t) - \gamma_X X_{\ell t} + \gamma_n n_{\ell t} \tag{1}$$

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.

At the beginning of each period, assembly plant quantities q_t and local characteristics X_t are revealed. Suppliers present then receive the resulting period profits π_ℓ. Next incumbent suppliers each privately draw a random scrap value ϕ_{it} which they claim by exiting the industry forever. All of n_{et} potential entrants privately draw random startup costs for each location, $\kappa_{i\ell t}$ for $i = 1, \ldots, L$. Potential entrants may set up permanently at a single location of their choosing or abstain and remain outside the industry forever. All potential entrants make their entry and site selection decisions and all incumbents make their exit decisions simultaneously.

A plant exits when it draws a scrap value larger than its discounted expected future profits, which include period profits and the discounted scrap value it eventually claims. Formulated recursively, the value of remaining in the industry is

$$V_\ell^{stay}(q, X, n) = E[\pi_\ell(q, X, n) + \beta V_\ell(q', X', n', \phi')$$

where $V_\ell(q, X, n, \phi) = \max\{\phi', \pi_\ell(q, X, n) + V_\ell^{stay}(q, X, n)\}$. Incumbents will chose to exit when $\phi > V_\ell^{stay}(q, X, n)$. Notationally, let $\chi_\ell(\phi, q, X, n)$ be the exit decision rule with $\chi_{it} = 1$ when supplier i chooses to exit and $\chi_{it} = 0$ when it chooses to stay.

Entry occurs at the location ℓ where the potential entrant receives the highest value after paying the entry cost. Let $\mu(q, X, n, \kappa)$ be the selection rule, and let $\mu_{it} = \ell$ if potential entrant i enters location ℓ and $\mu_{it} = 0$ if potential entrant i decides not to enter.

In this model a Markov perfect equilibrium is a policy function for potential entrants $\mu(q, X, n, \kappa)$, a policy function for incumbent suppliers in each location $\chi_\ell(q, X, n, \phi)$, transition probabilities $g(q', X', \phi', q, X, \phi)$, and a value function for each location $V_\ell(q, X, n)$ such that (1) given (q, X, n, ϕ, g), χ_l solves

$$V_\ell(q, X, n, \phi) = \begin{cases} \int \left(\pi_\ell(q', X', n') + V_\ell^{stay}(q', X', n', \phi')\right) g(q', X', \phi', q, X, \phi) dq' dX' d\phi' & : \chi_\ell = 0 \\ \phi & : \chi_\ell = 1 \end{cases}$$

(2) given (q, X, n, κ), $\mu(q, X, n, \kappa)$ maximizes $V_\mu^{stay}(q, X, n) + \kappa_\mu$ (or just κ_0 for $\sigma = 0$), (3) $n_\ell' = n_\ell -$

$\sum_{i=1}^{n_\ell}\chi_\ell(q,X,n,\phi_i)+\sum_j I[\mu(q,X,n,\kappa_j)=\ell]$, and (4) expected transition $g(q',X',\phi',q,X,\phi)$ frequencies match those observed.

Simplifications

The state space for the full game is unmanageably large for more than a few locations. In the present application, period profits depend on production quantities at over 100 assembly plants and the number of competitors in each of more than 1000 locations. Further, expectations on the number of competitors in each county depend on that location's changing characteristics. Tractability will come as the result of two adjustments to the model. The first simplification will be assumptions on the nature of competition that transform the problem into a collection of single-player dynamic discrete choice problems. The second is to summarize the thousands of state variables with two sufficient state variables.

First assume suppliers take competitor's locations as given. They do not hope to influence other suppliers with their exit decisions, and their expectations of the number of competitors do not depend on equilibrium or any equilibrium selection rule. Rather the transition matrix g matches the observed transitions, and agents take this transition matrix as given. The timing of the game is similar to that in Kalouptsidi (2011) and allows for exit rules to be based on the period's public observed number of suppliers rather than a number determined by equilibrium.

Suppliers are not directly concerned with the location of every competitor or the production quantities at each assembly plant. Instead they care about competitors and assemblers only inasmuch as competitors and assemblers affect their quantity of orders. The quantity $Q_{\ell t}=\sum_{a=1}^{A}q_a s_{\ell a}(n_t,X_t)$ summarizes competitors locations and assembly quantities in the exact way they enter suppliers' profit functions, so let suppliers track $\sum_{a=1}^{A}q_a s_{\ell a}(n_t,X_t,D)$ as a sufficient state variable. Likewise, suppliers are only concerned with characteristics like local wage inasmuch as they affect costs. Let $C_{\ell t}=C(X_{\ell t},n_{\ell t})$ be a state variable suppliers over which suppliers have expectations. Different market share functions were the only reasons π and V needed to be indexed by location, but since the sufficient state variable Q now tracks that, all locations have the same value function $V(C,Q)$.

5 Estimation

For estimation, periods will be five years and locations will be counties. Counties vary widely in size, so let each potential entrant make an extra entry cost draw for every 10,000 residents in a county. The number of potential entrants is set at 600 each period. County characteristics, the components of X, include interstate presence (interstate), local manufacturing wage (mfgwage), state manufacturing unionization rate (mfgunion), and population density (popden). In the calculation of market shares, and therefore Q, distances are weighted by the coefficient $\delta_{distance}$ and interstates are weighted by the coefficient $\delta_{interstate}$.

The only role of competition among suppliers in this model is to determine $Q_{\ell t}$ at each location. By assuming that evolution of $Q_{\ell t}$ is independent of any one plant's actions, the full dynamic model reduces from a dynamic game into a collection of single-agent dynamic decisions. The approach here is close in motivation to the oblivious equilibrium of Weintraub, Benkard, and Roy (2008).

Both the entry and exit decisions need to have choice specific logit error terms. Let the random scrap value be of the form $\phi + \epsilon^{exit}$ with ϕ a constant and ϵ^{exit} a random variable drawn from a type I extreme value distribution. Let another Type I extreme value random variable ϵ^{stay} be added to the incumbent's choice specific value function for not exiting. The value function now becomes $V(C, Q) = max\{\phi + \epsilon^{exit}, V^{stay}(C, Q) + \epsilon^{enter}\}$. Also, let the entry costs for each location be of the form $\sigma(\kappa + \epsilon_\ell)$ with κ and σ constants to be estimated and each ϵ_ℓ drawn from a type I extreme value distribution.

Estimations of the location-specific decisions follows a maximum likelihood approach. For each candidate set of parameter values, the sufficient state variables must be computed. Once state variables are computed, estimation follows the design pioneered by Rust (1987). The state variables are discretized into a eight by nine grid. The value function for incumbent suppliers is approximated by iteration. Rust (1987) parameterizes state variable transition matrices; the estimation here matches the transition frequencies exactly with the observed data, but the observed values of Q and C depend on the parameters δ and γ. The value function iteration is based on the Bellman equation

$$V(C, Q) = \ln(\exp\{\sum \beta(\gamma_p Q' - C' + V(C', Q'))G_Q(Q', C, Q)G_C(C', C, Q)\} + \exp\{\phi\}),$$

where G_Q and G_C are the observed transition probabilities.

Since the model has an additive logit error, the probability of an incumbent exiting location ℓ with state (C, Q) is $\rho_{x\ell} = \frac{exp\{\phi\}}{exp\{\phi\}+exp\{V^{stay}(C_\ell, Q_\ell)\}}$. The probability of an entrant establishing a plant in location ℓ is $\rho_{e\ell} = \frac{\exp V^{stay}(Q_\ell, C_\ell) - \kappa}{1+\sum \exp V^{stay}(Q_l, C_l)}$. The total log-likelihood function:

$$lnL(X, Q, D, \gamma, \delta) = \ln \sum_l \left\{ \rho_{e\ell} \sum \mu_\ell + \sum_{i=1}^{n_\ell} (\rho_{x\ell}\chi_i + (1 - \rho_{x\ell})(1 - \chi_i)) \right\}$$

This function is maximized through a simplex optimization routine.

6 Results

Table 7 reports the estimation results for three specifications of the dynamic model with the discount factor set as $\beta = (0.95)^5 \approx 0.773$). The negative coefficent on $\delta_{distance}$ indicates that assemblers are more likely to select nearby suppliers, so entrants have incentive to locate near supplier plants. Two of the specifications indicate the presense of agglomeration effects, as the γ_n coefficent is positive. The model does not identify whether the agglomeration effects are from productivity spillovers, labor pooling, or any of the other theorized agglomeration mechanisms, but it does seperate them from transportation costs and location natural advantages. Much of the cross-section variation in total costs C come from the natural avantages variables of wage and unionization rates.

Population density is highly correlated with the unionization rates and the number of suppliers in a county. Exclusion of the population density variable flips the sign on those variables. Inclusion seems the right approach, as population density is a measure of land rents and operating costs. The counterfactual estimates reported in the next section will use the specification 3 of the dynamic model.

If suppliers ignore profits in future periods (that is set $\beta = 0$), an incumbent's value function is just the maximum of their period profits and the scrap value. The existing literature of parts supplier site selection is one of static entries models closely related to my myopic case.[4] Table 8 reports estimates of this myopic case.

A supplier that enters into a location with low labor costs that is far from an assembly plant would appear in a static model not to care about transportation costs. If, however, the supplier expects an assembler to soon enter, it might still locate there even if it is sensitive to transportation costs. Because of the industry's migration, current distances to assemblers and densities of other suppliers are not good proxies for expectations for future distances and densities. This is exactly the bias observed in the myopic case. The coefficent γ_p, which is far lower (and negative in one case specification) in the myopic case. Because γ_p multiplies the shares determined by distance, the near zero γ_p makes suppliers appear not to care about distance. The dynamic model corrects this and implies distances to assembly plants are considered.

7 Counterfactual

Motor vehicle assemblers add assembly plants through an involved process near the end of which a short list of finalist sites are determined and the state and local governments for those sites offer subsidy packages.

[4]In the industry, the entry decision has received more attention than the plant closure decision. Most earlier works were estimated of a single cross-section, but the models have the similar logit structure and use many of the same location charac-terstics. (Smith and Florida 1994) and Head, Ries, and Swenson (1995) are early examples. The more recent work of Klier and McMillen (2008) uses spatial logits to correct for correlated error term, but still uses the static entry model setup.

This section examines a recent assembly plant placements and estimates how many additional suppliers the losing finalists would have added if their bids had won instead of lost.

For comparasion, the model with the actual assembler location is simulated. Next the counterfactual simulations exogenous move the assembly in question of its actual site to the losing finalist site. The number of suppliers and therefore the local costs will be adjust endogenously. Wages, population density, state transition matrices, and the production quantities at all other assembly sites are held constant.

On March 1, 2002 Hyundai announced the choice of Montgomery, Alabama over Glendale, Kentucky for its first North American assembly plant. The plant opened in 2005 and quickly began high production numbers. The plant appears in the 2006 data with an annual production quantity of over 230,000 unit. The plant begins affecting simulated entry decision in that period, and so the supplier counts for the simulated model and counterfactual model differ in 2011. That difference of 0.4 plants in Kentucky is the gains the model predicts would have come by 2011 from Kentucky outbidding Alabama for the Hyundai plant. Differences persist in entry decision in period beginning in 2011, when Kentucky has 6.7 more entrants in the counterfactual simulation than in the model simulation. The differences, however, are dwarfed in the inter-period volatility within simulation. Supplier counts for both simulations are reported in table 9, new entrant counts in table 10.

8 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 underestimate the role distance to customers and transportation costs play in clustering.

A model that can estimate the entry and closure decisions of parts suppliers can be used to estimate the benefits of attracting assembly plants. In the case of the Hyundai plant, the supplier plants in Kentucky would initailly be modest despite the large size of the assembly plant and expensive subsidy bids some states offered.

A Data Appendix

A.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.

A.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 dataset based on Dun & Bradstreet data from 1992 to 2006 employment measures are higher than in the QCEW or the Current Employment Statistics (CES), but by county-industry are highly correlated to both the QCEW and CES.

Table A.2 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 A.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.

References

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.

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.

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

Kalouptsidi, M. (2011): "Time to Build and Fluctuations in Bulk Shipping," Working paper, Princeton University.

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.

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.

Rubenstein, J. M. (2008): *Who Really Made Your Car? Restructing and Geographic Change in the Auto Industry*. W.E. Upjohn Institute for Employment Research, Kalamazoo.

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

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.

Thomadsen, R. (2005): "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," *The RAND Journal of Economics*, 36(4), 908–929.

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

Weintraub, G. Y., C. L. Benkard, and B. V. Roy (2008): "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 76(6), 1375–1411.

Table 1: Average Local Characteristics

Variable	Mean for incumbents in 1986	Mean for new entrants from 2006 to 2010
Latitude	40.26	39.50
County population	714745	597206
Incumbent suppliers in county	11.6	20.1
Assembler proximity	706.7	943.9
Interstate in county	0.800	0.836
UAW units in county	12.9	10.2
UAW units within 100 km	82.7	84.1
Observations	730	212

Table 2: Assembler output (in thousands)

	1986	1991	1996	2001	2006	2009
Northern Auto Alley	5816.7	4526.8	5182.6	4636.1	3960.7	1802.9
N Illinois	532.5	396.0	517.8	411.8	377.8	156.4
N Indiana	0.5	166.3	231.0	241.6	207.3	179.1
Michigan	3373.1	2624.5	2869.3	2687.5	2177.4	1146.1
N Ohio	1495.2	1204.7	1306.1	1046.7	983.1	319.9
Wisconsin	415.5	135.2	258.4	248.3	215.1	1.4
Southern Auto Alley	1365.5	2029.1	3026.5	3530.4	4367.2	2425.0
Alabama	0.0	0.0	0.0	84.5	698.1	467.8
Georgia	578.0	301.8	314.7	442.0	303.0	15.5
S Illinois	0.0	151.8	193.0	193.4	92.5	18.5
S Indiana	0.0	102.1	194.9	355.8	434.6	287.1
Kentucky	443.2	731.1	911.7	1141.6	1069.9	581.1
Mississippi	0.0	0.0	0.0	0.0	278.5	181.4
S Ohio	189.1	439.8	634.4	692.4	686.7	463.6
South Carolina	0.0	0.0	50.3	119.3	104.6	121.7
Tennessee	155.3	302.5	727.7	501.4	699.4	288.3
US outside Auto Alley	4080.6	2714.5	3232.8	3083.9	2461.3	1245.0
Canada	1832.7	1892.1	2368.1	2503.6	2497.4	1479.2
Mexico	263.3	960.9	1204.3	1829.5	1964.0	1506.6

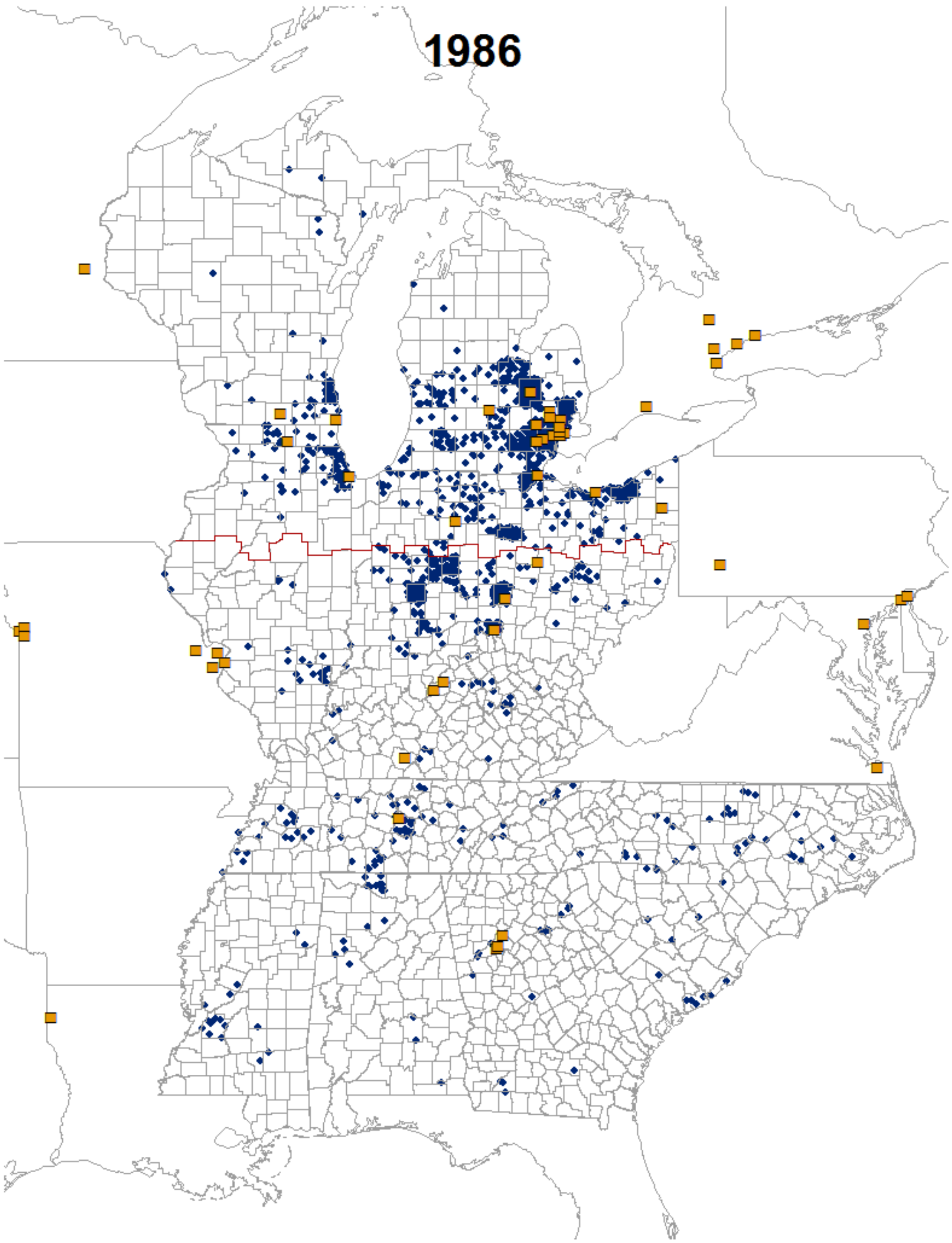

Figure 1: Supplier employment in auto alley and assembly plants in 1986. Squares indicate assembly plant locations. Each dot represents employment of 200 at supplier plants.

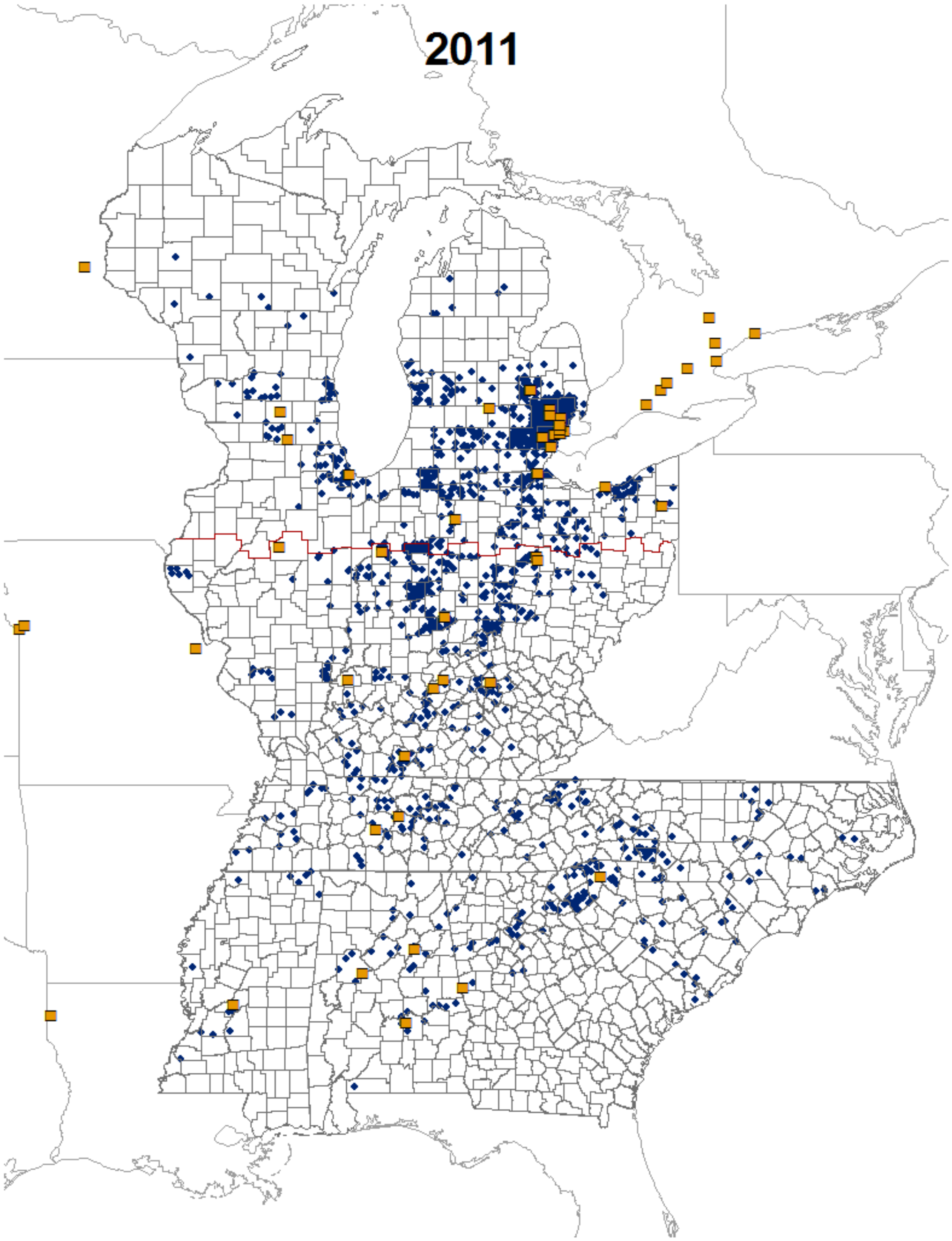

Figure 2: Supplier employment in auto alley and assembly plants in 2011. Squares indicate assembly plant locations. Each dot represents employment of 200 at supplier plants.

Table 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

Table 4: 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

Table 5: Supplier entry counts

	New entrant count in				
	1991	1996	2001	2006	2011
Northern Auto Alley	299	287	258	163	115
N Illinois	32	33	27	14	9
N Indiana	30	44	25	12	10
Michigan	156	150	160	114	68
N Ohio	61	44	35	15	14
Wisconsin	20	16	11	8	14
Southern Auto Alley	212	174	182	141	100
Alabama	11	9	7	14	17
Georgia	20	14	13	6	11
S Illinois	11	3	10	5	3
S Indiana	18	23	34	20	10
Kentucky	20	24	22	23	10
Mississippi	13	10	6	10	7
North Carolina	31	23	13	17	9
S Ohio	35	23	21	13	10
South Carolina	15	16	24	15	11
Tennessee	38	29	32	18	12

Table 6: 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 7: Dynamic model estimates ($\beta = (0.95)^5$)

	(1)	(2)	(3)
γ_p	0.00047	0.00080	0.00099
$\delta_{distance}$	-0.00014	-0.00033	-0.00024
$\delta_{interstate}$	0.0616	-0.0045	
$\gamma_{interstate}$			0.508
$\gamma_{mfgunion}$	0.25	-1.41	-0.61
$\gamma_{mfgwage}$	-0.0079	-0.0179	-0.0159
γ_{popden}		-0.00044	-0.00083
γ_n	-0.087	0.150	0.107
γ_{c1986}	4.39	-0.52	2.78
γ_{c1991}	-8.36	0.16	-1.91
γ_{c1996}	-0.12	-0.70	-2.82
γ_{c2001}	-0.47	4.67	-6.79
γ_{c2006}	1.93	-0.80	2.33
γ_{c2011}	-3.25	2.33	2.20
ϕ	-7.50	0.90	-0.30
κ	0.01	2.47	2.86

Table 8: Myopic model estimates ($\beta = 0$)

	(1)	(2)	(3)
γ_p	0.00004	-0.00092	0.00009
$\delta_{distance}$	-0.00550	-0.00085	-0.00037
$\delta_{interstate}$	0.0237	-0.0288	
$\gamma_{interstate}$			0.463
$\gamma_{mfgunion}$	0.95	-1.39	-0.70
$\gamma_{mfgwage}$	-0.0052	-0.0096	-0.0040
γ_{popden}		-0.00049	-0.00043
γ_n	0.046	0.195	0.087
γ_{c1986}	0.308	1.899	-0.821
γ_{c1991}	0.790	6.752	1.514
γ_{c1996}	3.138	-2.089	0.991
γ_{c2001}	0.673	6.279	3.081
γ_{c2006}	4.303	3.424	-3.560
γ_{c2011}	4.499	0.333	1.588
ϕ	-11.321	2.794	-1.514
κ	-0.029	-12.589	0.727

Table 9: Counterfactual Experiment: Supplier counts

	Suppliers in Kentucky		
Year	Actual	Model	Counterfactual
1986	21	21	21
1991	33	9.964	9.964
1996	47	11.407	11.407
2001	60	64.57	64.57
2006	69	131.932	131.932
2011	60	24.928	25.368

Table 10: Counterfactual Experiment: New entrants

	Entrants in Kentucy		
Period beginning	Actual	Model	Counterfactual
1986	20	7.163	7.163
1991	24	7.299	7.299
1996	22	62.398	62.398
2001	23	98.798	98.798
2006	10	9.245	9.249
2011		64.615	71.323