NET Institute*

www.NETinst.org

Working Paper #08-14

September 2008

Network Effects and Geographic Concentration of Industry

Zhu Wang
Federal Reserve Bank of Kansas City.

Yi Xu
New York University

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Network Effects and Geographic Concentration of Industry*

Zhu Wang and Yi Xu†

September 30, 2008

Abstract

This paper provides a theory of "family network," in contrast to "local externalities," to explain the geographic concentration of industry. For many industries, one most important source of entrants is spinoffs, who typically locate near parent firms and benefit from knowledge linkage and business relation within the family network. As a result, firms are more likely to enter and less likely to exit if they are associated with a large family. Using a unique dataset of US automobile industry in its early years, we identify six historically important production centers and sixty spinoff families. Our empirical analysis disentangles the effect of "family networks" from other "local externalities," and provides strong evidence that it was the former rather than the latter that caused the geographic concentration of US automobile production.

JEL classification: J6; L0; R1

Keywords: Spinoff; Entry and Exit; Geography of Industry

---

*Preliminary draft, please do not cite or circulate without authors' permission. We thank Boyan Jovanovic and Robert Lucas for helpful discussions, and Christian Hung for excellent research assistance. Xu gratefully acknowledges financial support from the NET Institute (www.netinst.org) and the Kauffman Foundation. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

†Wang: Economic Research Department, Federal Reserve Bank of Kansas City. Email: zhu.wang@kc.frb.org. Xu: Department of Economics, New York University. Email: daniel.xu@nyu.edu.

1 Introduction

Geographic concentrations of individual industries are striking features of the economic landscape. The agglomeration of high-tech industries in Silicon Valley and automobile industry in Detroit are two classic examples. Many recent studies have attempted to investigate why agglomerations took place, and hope to shed light on the nature of increasing-return technologies and spillovers that are thought to be important driving forces of endogenous growth and international trade.

Most existing literature (e.g., Fujita et. al. 1999; Quigley 1998, Krugman 1997, Henderson 1988) suggests that agglomerations are the result of positive local spillovers. While individual firms are perfectly competitive and subject to constant returns to scale, the agglomeration of economic activity generates externalities that raise the productivity of all firms in a particular industry that share a given geographic location. These externalities are assumed and their sources are not specified. Proposed mechanisms, following Marshall (1920), include knowledge spillovers, joint development and use of human capital, the attraction of suppliers and service firms to the region, and cooperation in R&D. Despite of the differences in details, all of these mechanisms lead to the fact that regions with already a large number of firms in a certain industry become attractive for further firms of this industry.

Meanwhile, congestion costs, associated with limited local supplies of housing or other non-traded goods or factors, work against agglomeration. For example, as agglomeration forms, the price of housing is bid up in the area. To attract workers, firms must compensate workers for the relatively high cost of living. The productivity of labor in agglomerated regions then has to justify these higher wages. Therefore, costs associated with congestion, act as a centrifugal force which prevents economic activity from becoming too agglomerated.

In this paper, we seek a better understanding of agglomerations by investigating the network effects within each family of spinoff firms. For many industries, one most important source of entrants is spinoffs, who work as employees for existing firms and later start their own business in the same industry. Spinoffs typically locate near the parents, and benefit from knowledge linkage and business relation within the family network. As a result, firms are more likely to enter and less likely to exit if they are associated with a large family. Therefore, geographic concentrations of individual industries could be driven mainly by this "family network effect," a special local externality within the family network.

Using a unique dataset of US automobile industry, we test the "family network effect" against other "local externalities" in terms of their contribution to the geographic agglomeration. The former is closely related with the spinoff process, and depends crucially on the heritage and network of each individual spinoff family. In contrast, the latter simply depends on the total number of firms in each location. Our empirical findings suggest that it was the "family network effect" rather than other "local externalities" that caused the geographic concentration of US automobile production. In fact, after we control for the "family network effect," the other "local externalities" show negative effects on agglomeration, which means congestion.

Related to our paper, Klepper (2007) investigates spinoffs and the evolution of Detroit as the capital of the US automobile industry. He explains the agglomeration in Detroit using a theory that disagreements lead employees to spin off from incumbent firms. Our paper differs from Klepper (2007) in several important aspects. First, we empirically disentangle the effects from local externalities, which is emphasized by the economic geography literature, and the effects from family networks through spinoffs. For this purpose, we use our data to identify six historically important automobile production centers to isolate location specific characteristics or spillovers. Second, we construct a theoretical model that depends heavily on the industry selection effect similar to Hopenhayn (1992), in which incumbents and potential spinoffs make decisions based on their quality and location-specifics.

It is noteworthy that our theory, like many other previous studies, explains why conglomeration form, but not necessarily predict where they will form. In our framework, all locations are ex ante identical such that it is not pre-determined that which regions will be home to industry clusters and which will not. In reality, the natural features of regions are likely to impact industry location. And we do find some evidence that the location specific effects play some roles (see also Ellison and Glaeser 1999).

The paper is organized as follows. In section 2, we provide a simple industry equilibrium model with spinoff entries. The model shows the geographic concentration of industry is mainly driven by the agglomeration of firms from major spinoff families. In section 3, we test our theory using a dataset of US automobile industry. The dataset is unique in the sense that it helps to uncover the heritage of each spinoff firm, which allows us to distinguish the "family network effect" from other "local externalities." Our findings suggest that the "family network effect" rather than other "local externalities" caused the geographic concentration of US automobile production. Section 4 concludes.

2 Model of Spinoffs and Family Network

In this section, we provide a simple industry equilibrium model with spinoff entries. We assume a potential spinoff entrant shares the same quality with all other firms in the family. Here the "family" is defined as all spinoff firms who share the same ancestor, including the ancestor who may not be a spinoff himself. Hence, all firms in the family are of the same quality. To the extent that this family-specific quality may result from knowledge sharing or business relation within the family network, we may regard it as a special form of "local externality".[1]

2.1 Individual Firm's Problem

Time is discrete and indexed by $t = 1, 2, 3, ..., \infty$. The model industry comprises of firms of different quality $s \in [0, \bar{s}]$ located at various production centers j. For simplicity, we assume

[1] Note our model may also be consistent with alternative interpretations without refering to externality, for example, spinoffs inherit ability from parents.

that a firm starting at center j will operate at the same location for the rest of its life.[2] The industry structure at each period t can be summarized by m_t. Each of its element $m_t(s,j)$ is the total mass of firms of quality s at location j.

At the beginning of period t, all incumbent firms engage in product market competition by taking industry price p as given. Each firm decides on the optimal quantity of output based on its quality and location characteristics. Their period profit is determined by $\pi(s;j,p)$, where j denotes the firm's location characteristics (not changing overtime), p is the industry price at period t.

Once incumbent firms obtain their profit, they decide whether to stay in this industry or to leave by taking some outside options ϕ^x . The distribution of these outside options is i.i.d. across firms. Thus we have the incumbent's problem defined as:

$$\begin{aligned} VC(s;j,\bar{p}) &= \beta \int V(s;j,\bar{p},\phi^{x\prime})dF(\phi^{x\prime}) \\ V(s;j,\bar{p},\phi^x) &= \pi(s;j,p) + max\{VC,\phi^x\} \end{aligned}$$

Notice here that we have been abstract in defining the sequence of the industry price $\bar{p}$. We will defer the discussion of this after we define the entrants' decisions.

The industry potential entrants at each production center make their entry decisions at the same time that incumbents are making their exit decisions. However, there are two different types of entrants. The first type, which we call De Novo entrants, are the entrepreneurs who have never worked in this industry. We assume that De Novo entrants don't know their type before they enter and they are randomly allocated to an location j with probability $\mu(s,j)$. Once they pay the fixed sunk cost ϕ^e, they have their initial draw of type s from distribution $\mu(s,j)$. Their total mass of entry M solves the free-entry condition

$$\int_s VC(s;j,\bar{p})d\mu(s,j) = \phi^e.$$

[2] This can be justified, for example, by the costs of relocation.

Finally, each incumbent also has a probability γ of giving birth to a potential spinoff each period. We assume the potential spinoff entrant shares the same quality s with all other firms in the family, and knows his quality while making his entry decision. If he doesn't choose to start a firm at current period, his opportunity is foregone once for all. These potential spinoffs will enter if their value of entry is higher their random outside option ϕ^x, i.e.

$$VC(s; j, \bar{p}) \geq \phi^x.$$

2.2 Transition of Industry Structure and Price

Let's first define the transition of the mass of firms of quality s in market j, with current industry price p. It depends on the number of exits, spinoffs, and De Novo entries at each state (s, j). Explicitly, we have

$$m_{t+1}(s, j) = m_t(s, j)(1 + \gamma)\chi_t + M\mu(s, j),$$

where $\chi_t = \Phi^x(VC(s; j, \bar{p}))$ is probability of staying, given the cumulative distribution Φ^x of the industry outside option.

Next, we will also need to define the industry output market. Aggregate demand is given by the inverse demand function $D^{-1}(Q)$. Industry price will satisfy each period such that

$$p = D^{-1}\{\int q(s, j, p)m_t(s, j)\}.$$

2.3 Industry Equilibrium

We define the industry equilibrium by a bounded sequence of p_t, m_t, χ_t, and M_t such that:

(1). χ_t solves incumbent firms and potential spinoffs' dynamic optimization problem.

(2). Potential De Novo entrants M_t satisfy zero profit condition.

(3). p_t clears product market each period.

(4). m_t is defined recursively given m_0, M_t, and χ_t

Following immediately Hopenhayn (1992), there exists a competitive equilibrium, where the path of industry price $\bar{p}$ is deterministic given initial industry structure m_0. Particularly, there exists a stationary equilibrium, defined as an output price $p^* \geq 0$, a mass of entrants M^*, a measure of incumbents $m^*(s, j)$, and policy function χ^*, such that for $p_t = p^*$, $m_t = m^*$, $M_t = M^*$, and $\chi_t = \chi^*$ is an equilibrium from $m_0 = m^*$.

2.4 Model Implications

The model has the following testable implications.

Proposition 1 *A potential spinoff firm is more likely to enter and less likely to exit if it belongs to a high quality family.*

Proof. Given $\pi(s; j, p)$ is strictly increasing in s, continuous, and bounded, standard dynamic programming arguments can be used to show that $VC(s; j, \bar{p})$ is continuous in s and for $\bar{p} > 0$ strictly increasing in s. Thus we know that for each period t, $\Phi^x(VC(s; j, \bar{p}))$ is strictly increasing in s, given the same production location j. ■

Proposition 2 *High-quality incumbents have higher probability of producing spinoff firms.*

Proof. This is straightforward because all incumbents have the same probability γ of having a potential spinoff, while the spinoff's probability of entering χ^* is increasing in s. ■

Proposition 3 *The higher the family quality, the bigger the family size.*

Proof. As shown above, higher quality incumbents produce more spinoffs on average. Meanwhile, the incumbents of a higher-quality family have lower probability of exit. ■

Proposition 4 *If there is positive entry and exit in the stationary equilibrium, spinoff firms have lower probability to exit than De Novo entrant, given the same location j.*

Proof. The stationary distribution is defined by $m^* = m^*(1+\gamma)\chi^* + M^*\mu$, so $m^* = \frac{M^*}{1-(1+\gamma)\chi^*}\mu$. The distribution of spinoff firms is $m^*\chi^* = \frac{M^*\chi^*}{1-(1+\gamma)\chi^*}\mu$. Since χ^* is strictly increasing in s, the distribution of the abilities of spinoff firms strictly dominates that of the De Novo entrant firms, which is μ. ■

3 Empirical Analysis

In this section, we estimate/test our model using a unique dataset of the US automobile industry. The dataset includes US companies that sold at least one automobile to the public during the first 75 years of the industry (1895-1969), a total of 780 firms with their entrepreneurial and geographic characteristics.

3.1 Data Sources

The data sources come from several industry references. First, Smith (1968) provides a list of every make of automobile produced commercially in the United States from 1895 through 1969.[3] The book lists the firm that manufactured each car make, the firm's location, the years that the car make was produced, and any reorganizations and ownership changes that the firm underwent. Smith's list of car makes was then used to derive entry, exit and geographic location of firms.[4]

Second, Kimes (1996) provides comprehensive information for every automobile make produced in the US from 1890 through 1942. Using Kimes (1996), we are able to collect additional biographical information about the entrepreneurs who founded and ran each individual firm. An entrepreneur was then categorized into one of the following three groups: De Alio entrants, Spinoff entrants and De Novo entrants. The first group includes entrepreneurs who had prior experience in related industries before starting an automobile firm. The second group includes entrepreneurs who had worked as employees in existing automobile firms before starting their

[3] The original book published in 1968 was updated to include information up to 1969.
[4] The entry and exit dates are based on the first and last year of commercial production.

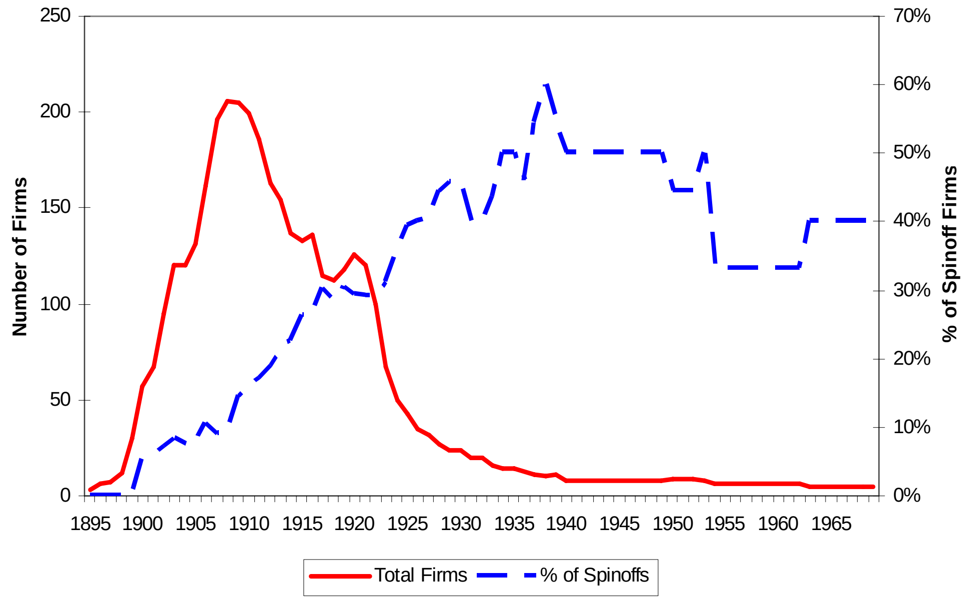


Figure 1: The Evolution of US Automobile Industry: 1895-1969

own. The last group includes those with no identifiable background information. Kime's information was then used to derive family linkage between individual firms, in other words, we construct family trees for each spinoff firm.

Third, Bailey (1971) provides a list of leading automobile makes from 1896-1970 based on top-15 annual sales. This allows us to identify top automobile producers during those periods.

3.2 Industry Overview

As shown in Figure 1, the automobile industry went through a tremendous development in its first 75 years, evolving from a small infant industry to a gigantic, concentrated, mature industry. During the process, an industry shakeout started around 1910, as the number of firms fell steadily from a peak of 206 in 1908 to 8 in 1942. Meanwhile, the percent of spinoff firms continued to increase, from almost zero in 1900 to 60% in 1940. Later on, the ratio of spinoffs started to decline but that was because some top parent firms (who are not spinoffs themselves) of major spinoff families outlived their children.

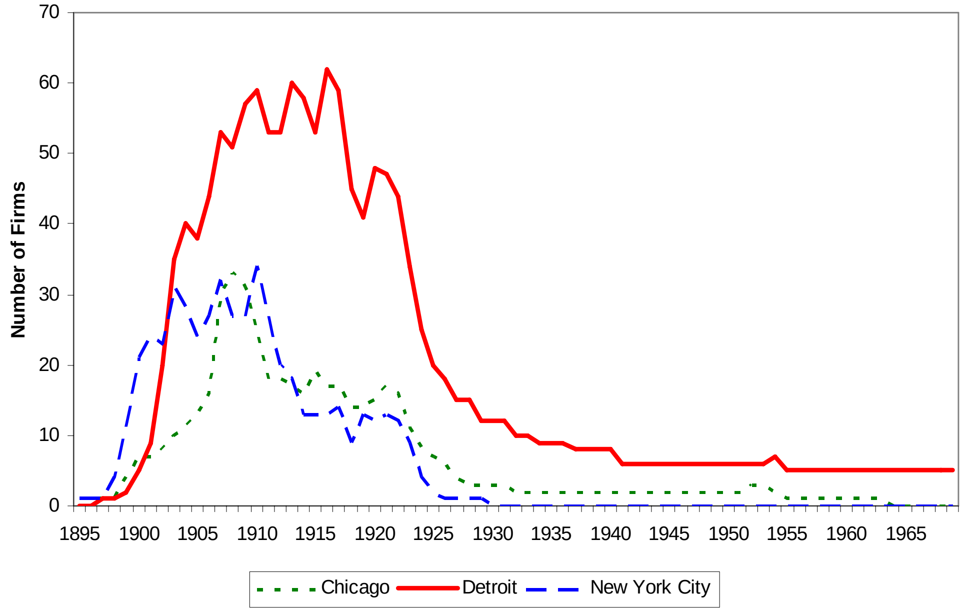


Figure 2: Geographic Dynamics of Automobile Firms

The industry also went through significant geographic dynamics over those years. Using the number of firms as the criterion, we identify six historically important automobile production centers, namely St. Louis, Chicago, Indianapolis, Detroit, Rochester and New York City.[5] As shown in Figure 2, the industry initially started in Chicago and New York City before 1900. Soon after, Detroit and other centers caught up quickly (See Figure A1 in the Appendix for maps of geographic dynamics of automobile production in the U.S. between 1900-1925). Until the shakeout started around 1910, the number of firms in every production center grew strongly. After that, the number of firms started to fall in most centers (except for Detroit, where the shakeout did not start until 1918) but at different rates. Eventually, the industry production was dominated by a few surviving Detroit firms.

The entry and exit pattern of automobile firms were also different across spinoff families. Using the family trees that we constructed, we identify a total of 197 firms associated with

[5] A city is counted as an automobile production center city if it had at least five automobile producers in year 1910. We then define the region within 100 miles of the center city as the production center, named after the center city.

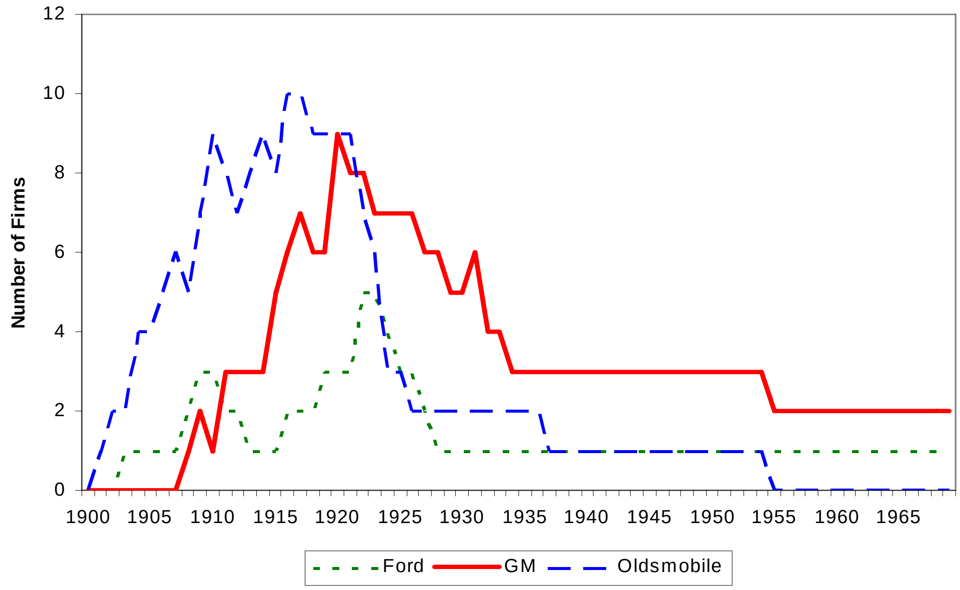


Figure 3: Family Dynamics of Automobile Firms

60 families. Among them, Ford, GM and Oldsmobile were the three biggest families, each generating 12-17 spinoffs (see Table A1 in the Appendix for the family trees). As expected, most spinoffs locate near the parents: (e.g., 76% of spinoffs locate near the parents in the top three families). Figure 3 presents the time path of firm numbers for the top three families. As shown, each family also went through a shakeout, though the exact patterns were not quite the same.

3.3 Regression Analysis

To summarize, our automobile dataset includes the following information:

(1) The entry year of each firm.

(2) The exit year of each firm.

(3) The type of each exit.

(4) The background of each entrepreneur: De Alio, Spinoffs or De Novo.

(5) The quality of firms in terms of producing top makes in the industry.

(6) The location of each firm.

(7) The six automobile production centers.

(8) The family linkage for each spinoff firm.

Using the above information, we create the following dummy variables used in our regressions (indexed by firm and year). Whenever needed, additional explanation is given in parentheses.

◦ Firm Death (The firm exited in the current period).

◦ Firm Birth (The firm gave birth to a spinoff in the current period).

◦ Production Center (Seven dummies corresponding to St. Louis, Chicago, Indianapolis, Detroit, Rochester, New York City and other places).

◦ De Alio (The firm was founded by an experienced entrant).

◦ Spinoff (The firm was founded by a spinoff entrant).

In addition, we create the following variables:

◦ Center Size (The number of firms in the production center where the firm is located).

◦ Center Top (The number of top firms in the production center where the firm is located).

◦ Family Size (The number of firms in the family to which the firm belongs).

◦ Family Top ((The number of top firms in the family to which the firm belongs).

◦ Firm Age.

◦ Year.

Tables A2 and A3 in the Appendix provide descriptive statistics of the regression variables, both at firm×year level and firm level. From Table 1, we can see that the firm death rate is about 17% per year and the average age of a firm is about 7 years. These are not unusual for a growing industry. Meanwhile, the firm spinoff (birth) rate is about 2% per year. We can also see 59% of firms are De Alio entrants, 20% are Spinoffs, and 20% firms ever made the top firm list. On average, a location has 36 firms (the range is from 1 to 96), among which 6 of them are top firms (the range is from 0 to 19); a spinoff family has 1.6 member firms (the range is from 1 to 10), and 0.5 firms are top ones (the range is from 0 to 6). From Table 2, we can see that about 18% of all firms are spinoffs from existing firms, and one half are experienced entrants. Among all firms, 6% of them ever made the top list.

In the following analysis, we ran logit regressions using firm-year observations with Firm Death or Firm Birth as the dependent variable. As suggested by our theory, to the extent that each firm's outside options ϕ^x is logistically distributed, our exercises are equivalent to estimating firms' policy function of entry and exit. The data range we use is from 1895-1942, including 776 firms and 4472 firm-year observations.[6]

3.3.1 Firm Death Analysis

Table 1 presents the regression results with Firm Death as the dependent variable. The findings support the implications of our model.

First, the variable "Family Size" has a negative coefficient as predicted by Propositions 1 and 3. The coefficient is always statistically significant at 0.1% level in various model specifications. The magnitude of the coefficient is also economically significant: The corresponding odds ratio implies that the relative death rate of firm will drop by 14% as the number of firms in a family increases by one. Given the fact that firms are of different size in reality, we also try "Family Top" as an alternative measure. Again, it has a negative coefficient as predicted by Propositions 1 and 3. The coefficient is also statistically significant at 0.1% level, and the odds ratio implies that the relative death rate of firm will drop by 36% as the number of top firms in a family increases by one.

Second, the variable "Spinoff" has a negative coefficient as predicted by Proposition 4. However, whether the coefficient is statistically significant or not depends on model specifications. Particularly, when we include "Family Top" instead of "Family Size" into the regression,

[6] Given the information provided in Kimes (1996), we collect biographical information about the entrepreneurs up to 1942, before the US entered the WWII.

Table 1. Logit Regressions on Firm Death Rates

Firm Death	(1)	(2)	(3)	(4)	(5)	(6)
De Alio			-0.479***	-0.474***	-0.515***	-0.506***
			(0.096)	(0.096)	(0.097)	(0.098)
Spinoff			-0.330**	-0.166	-0.340**	-0.182
			(0.146)	(0.139)	(0.150)	(0.143)
St. Louis	0.015	0.292	0.415	0.287	0.133	0.179
	(0.239)	(0.271)	(0.274)	(0.258)	(0.317)	(0.271)
Chicago	-0.107	0.099	0.108	-0.009	-0.075	-0.050
	(0.140)	(0.169)	(0.171)	(0.153)	(0.204)	(0.161)
Indianapolis	-0.560***	-0.385**	-0.321*	-0.489***	-0.516**	-0.561***
	(0.155)	(0.175)	(0.176)	(0.164)	(0.203)	(0.173)
Detroit	-0.422***	-0.381***	-0.245**	-0.643**	-0.295**	-0.462
	(0.110)	(0.112)	(0.116)	(0.268)	(0.120)	(0.309)
Rochester	-0.253	0.007	0.041	-0.162	-0.192	-0.207
	(0.196)	(0.230)	(0.231)	(0.202)	(0.263)	(0.206)
New York	0.213*	0.418***	0.432***	0.261**	0.278	0.261*
	(0.127)	(0.158)	(0.160)	(0.131)	(0.185)	(0.134)
Center Size		0.006**	0.006**		0.001	
		(0.003)	(0.003)		(0.004)	
Family Size			-0.140***		-0.149***	
			(0.040)		(0.042)	
Center Top				0.043*		0.024
				(0.023)		(0.027)
Family Top				-0.439***		-0.441***
				(0.069)		(0.070)
Firm Age	-0.054***	-0.053***	-0.050***	-0.038***	-0.051***	-0.042***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
Year	0.026***	0.031***	0.037***	0.031***		
	(0.006)	(0.007)	(0.007)	(0.007)		
Constant	51.462***	61.627***	71.370***	60.502***	-1.472***	-1.610***
	(12.086)	(13.188)	(13.881)	(13.309)	(0.420)	(0.397)
Year Dummies					Y	Y
Observations	4458	4458	4458	4458	4364	4364

Note: Standard errors are in parentheses under coefficient values. One, two and three * indicate statistical significance at the 10, 5 and 1% levels, respectively.

the coefficient of “Spinoff” is no longer significant. One interpretation is that the “Family Top” is a better measure of the quality of the family so that it captures the difference between Spinoff and De Novo firms.

Third, the variable “Center Size” has a positive coefficient, and it is statistically significant. This suggests that there is no local positive externality, but rather local congestions. We also try “Local Top” as an alternative measure. As expected, it has a larger positive coefficient than that of “Center Size,” which implies top firms in a location cause bigger congestions.

Fourth, among all six location dummies, three are statistically significant (Detroit, Indianapolis and New York). Particularly, the two dummies “Detroit” and “Indianapolis” have negative coefficients, which suggest some location-specific advantage. In contrast, “New York” has a positive coefficient, which suggests some location-specific disadvantage.

Finally, the variables “De Alio” and “Firm Age” both have negative coefficients and statistically significant. This is consistent with the explanation that firm age and experience indicate its quality. Meanwhile, we use the variable “Year” to capture the changing threshold of surviving in the industry. As expected, the coefficient is positive and statistically significant.

3.3.2 Firm Birth Analysis

Table 2 presents the regression results with Firm Birth as the dependent variable. The findings also support the implications of our model.

First, the variable “Family Size” has a positive coefficient as predicted by Propositions 1 and 3. The coefficient is always statistically significant at 0.1% level in various model specifications. The magnitude of the coefficient is economically significant: The corresponding odds ratio implies that the relative birth rate of firm will increase by 40% as the number of firms in a family increases by one. We also try “Family Top” as an alternative measure. The result is similar.

Second, without controlling for “family network effect,” the variable “Center Size” has a positive coefficient and it is statistically significant. This seems to suggest local positive

Table 2. Logit Regressions on Firm Birth Rates

Firm Birth	(1)	(2)	(3)	(4)	(5)	(6)
St. Louis	-0.144	1.038	0.294	-0.015	-1.135	-0.797
	(1.045)	(1.160)	(1.155)	(1.069)	(1.220)	(1.084)
Chicago	0.620	1.529**	0.996*	0.699	-0.019	0.241
	(0.461)	(0.607)	(0.598)	(0.483)	(0.651)	(0.497)
Indianapolis	0.306	1.088*	0.736	0.457	-0.286	-0.122
	(0.481)	(0.593)	(0.579)	(0.492)	(0.634)	(0.516)
Detroit	1.299***	1.575***	0.673*	0.265	0.415	2.147***
	(0.334)	(0.365)	(0.395)	(0.615)	(0.384)	(0.763)
Rochester	0.752	1.815***	1.264*	0.832	0.067	0.511
	(0.513)	(0.690)	(0.674)	(0.523)	(0.764)	(0.534)
New York	0.883**	1.758***	1.195**	0.853**	0.332	0.788*
	(0.413)	(0.564)	(0.544)	(0.418)	(0.603)	(0.422)
Center Size		0.022***	0.012		-0.015	
		(0.009)	(0.009)		(0.013)	
Family Size			0.333***		0.338***	
			(0.044)		(0.047)	
Center Top				0.056		-0.125*
				(0.049)		(0.064)
Family Top				0.324***		0.341***
				(0.067)		(0.069)
Firm Age	0.029	0.039*	0.057**	0.033	0.072***	0.055**
	(0.020)	(0.021)	(0.022)	(0.021)	(0.023)	(0.022)
Year	-0.027	-0.016	-0.051**	-0.039**		
	(0.019)	(0.021)	(0.022)	(0.019)		
Constant	47.347	24.187	91.663**	69.611*	4.557***	-4.828***
	(35.527)	(40.241)	(42.260)	(36.980)	(1.143)	(1.056)
Year Dummies					Y	Y
Observations	4472	4472	4472	4472	4121	4121

Note: Standard errors are in parentheses under coefficient values. One, two and three * indicate statistical significance at the 10, 5 and 1% levels, respectively.

externalities. However, after we control for “Family Size” or “Family Top”, the coefficient of “Center Size” or “Center Top” lost its statistical significance. Moreover, when we introduce year dummies, their coefficients turn negative, which again suggests local congestions.

Third, three location dummies (Detroit, Rochester and New York) show positive signs, which suggest some advantage of encouraging spinoff entries. However, these coefficients are not always statistically significant across different model specifications.

Finally, the variable “Firm Age” has positive coefficient and statistically significant. This is consistent with the explanation that firm age indicates its quality. Meanwhile, we use the variable “Year” to capture the changing threshold of entering the industry. As excepted, the coefficient is positive and statistically significant.

4 Conclusion

This paper provides a theory of “family network” to explain the geographic concentration of industry. In contrast to the traditional view that spatial agglomeration of industry is caused by “local externalities”, we find that the “family network effect” could actually been the driving force.

For many industries, one most important source of entrants is spinoffs, who work as employees for existing firms and later start their own business in the same industry. Spinoffs typically locate near the parents, and benefit from knowledge linkage and business relation within the family network. As a result, firms are more likely to enter and less likely to exit if they are associated with a large family.

Using a unique dataset of US automobile industry in its early years, we identify six historically important production centers and sixty spinoff families. Our empirical analysis then disentangles the effect of “family networks” from other “local externalities.” In fact, after we control for the “family network effect,” other “local externalities” show negative effects on firm entry and survival, which means local congestions. This provides strong evidence that it was the former rather than the latter that caused the geographic concentration of US automobile production. Although we do not explicitly specify in our analysis the source of the "family network effect" and whether it has to be an externality or not, we have clearly shown that family-specific effects play a central role in the geographic concentration of industry.

References

[1] Bailey, L.S. (1971). *The American Car since 1775.* Automobile Quarterly: New York.

[2] Dumais, G., G. Ellison and E. L. Glaeser, (2002). "Geographic Concentration As A Dynamic Process," *Review of Economics and Statistics*, 84(2): 193-204.

[3] Ellison, G. and E. L. Glaeser, (1999). "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?," *American Economic Review*, 89(2): 311-316,

[4] Ellison, G. and E.L. Glaeser. (1997). "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy,* 105: 889-927.

[5] Fujita, M., P. Krugman, and A. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

[6] Franco, April, and Darren Filson (2006). "Spin-Outs: Knowledge Diffusion Through Employee Mobility," *Rand Journal of Economics* 37, 841-860.

[7] Jovanovic, Boyan (1982). "Selection and Evolution of Industry," *Econometrica* 50, 649-670.

[8] Hanson, Gordon (2001). "Scale Economies and the Geographic Concentration of Industry," *Journal of Economic Geography*, 1(3): 255-276.

[9] Henderson, J.V. (1988). *Urban Development: Theory, Fact, and Illusion.* New York: Oxford University Press.

[10] Hopenhayn, Hugo (1992). "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5): 1127-50.

[11] Klepper, Steven (2007). "Disagreements, Spinoffs, and the Evolution of Detroit as the Capital of the U.S. Automobile Industry," *Management Science*, 53 (4): 616-631.

[12] Klepper, Steven (2002). "The capabilities of new firms and the evolution of the US auto-

mobile industry," *Industrial and Corporate Change*, vol. 11(4), 645-666.

[13] Kimes, Beverly R. (1996) *Standard Catalog of American Cars, 1890-1942*, third edition, Iola, WI: Krause Publications Inc.

[14] Krugman, P. (1997). *Development, Geography, and Economic Theory.* Cambridge: MIT Press.

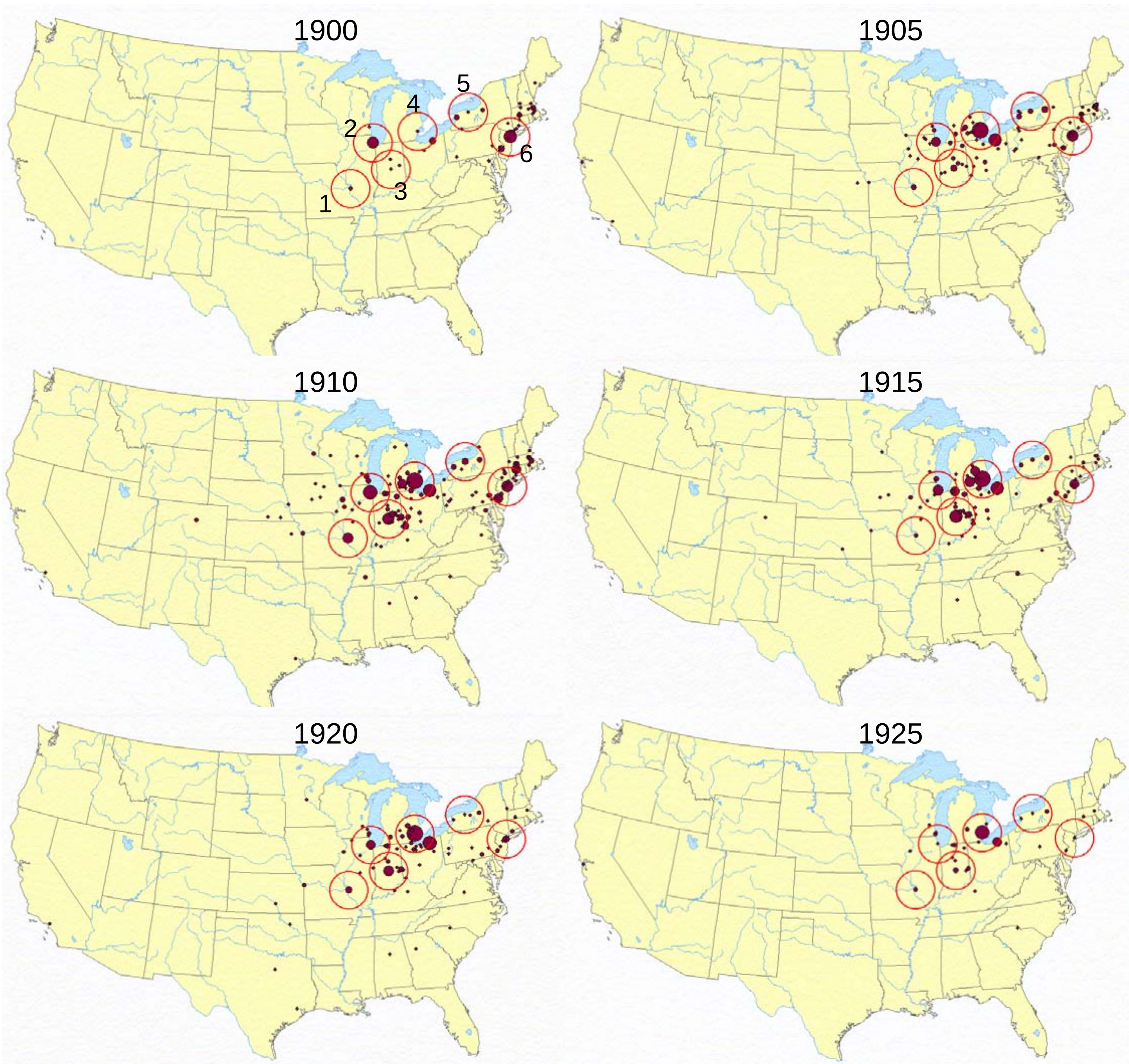
[15] Krugman, P. (1991). "Increasing Returns and Economic Geography." *Journal of Political Economy* 99: 483-499.

[16] Marshall, A. (1920). *Principles of Economics.* New York: Macmillan.

[17] Quigley, J.M. (1998). "Urban Diversity and Economic Growth." *Journal of Economic Perspectives,* 12: 127-138.

[18] Smith, Philip H. (1968). *Wheels within Wheels*, New York: Funk and Wagnalls.

Figure A1. Geographic Dynamics of US Auto Production Centers (1900-1925)



Firms in a City

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8-10
- 11-15
- 16+

Production Centers (red circles):

1. Saint Louis
2. Chicago
3. Indianapolis
4. Detroit
5. Rochester
6. New York

Table A1. Spinoff Trees of Top US Auto Producer Families

General Motors (1908 -)	Chevrolet (1911 - 1918)	Gardner (1920 - 1931)	
		Monroe (1914 - 1920)	
	Durant (1921 - 1931)	DeVaux (1931 - 1933)	
		Flint (1902 - 1903)	
	Nash (1917 - 1954)	Ajax (1925 - 1926)	
		American Motors Corp. (1954 -)	
		Lafeyette (1920 - 1924)	
	Farmack (1915 - 1916)	Drexel (1917 - 1917)	
	Chrysler (1924 -)		
	Daniels (1915 - 1923)		
	Lincoln (1920 - 1922)		
	Little (1911 - 1913)		
	Olympian (1916 - 1920)		
	Paige-Detroit (1919 - 1928)		
	Welch-Detroit (1909 - 1909)		
Oldsmobile (1901 - 1917)	Maxwell (1904 - 1925)	Briscoe (1914 - 1921)	Earl (1921 - 1923)
			Edwards-Knight (1912 - 1913)
		Motorette (1910 - 1911)	
		Spartan (1910 - 1910)	
	Chalmers (1906 - 1921)	Saxon (1913 - 1922)	Liberty (1916 - 1923)
	Northern (1902 - 1908)	King (1911 - 1923)	
	Hudson (1909 - 1954)		
	Neilson (1907 - 1907)		
	Nelson (1916 - 1920)		
	Owen (1909 - 1910)		
	Reo (1904 - 1936)		
	Warren-Detroit (1909 - 1914)		
Ford (1903 -)	E-M-F (1908 - 1910)	Columbia (1916 - 1924)	
		Everitt (1909 - 1912)	
		Hanson (1919 - 1923)	
		Rickenbacker (1922 - 1926)	
	Hupmobile (1909 - 1940)	Emerson (1917 - 1918)	Campbell (1918 - 1919)
		Hupp-Yates (1910 - 1916)	
		R.C.H. (1911 - 1916)	
	Gray (1922 - 1926)		
	Wills St. Claire (1922 - 1927)		

Table A2. Data Summary Statistics (Firm×Year Level)

Variable	Obs	Mean	Std. Dev.	Min	Max
Firm Death	4472	0.17	0.38	0	1
Firm Birth	4472	0.02	0.14	0	1
De Alio	4472	0.59	0.49	0	1
Spinoff	4472	0.20	0.40	0	1
Top Firm	4472	0.20	0.40	0	1
Center Size	4472	35.81	23.66	1	96
Family Size	4472	1.56	1.59	1	10
Center Top	4472	6.04	5.64	0	19
Family Top	4472	0.50	1.16	0	6
Firm Age	4472	6.85	7.18	1	43
Year	4472	1913	8	1895	1942

Table A3. Data Summary Statistics (Firm Level)

Variable	Obs	Mean	Std. Dev.	Min	Max
De Alio	776	0.52	0.50	0	1
Spinoff	776	0.18	0.38	0	1
Top Firm	776	0.06	0.24	0	1
Entry Year	776	1908	6	1895	1939