Working Paper Series

Competitors, Complementors, Parents and Places: Explaining Regional Agglomeration in the U.S. Auto Industry

WP 13-04

Luís Cabral
New York University and CEPR

Zhu Wang
Federal Reserve Bank of Richmond

Daniel Yi Xu
Duke University and NBER

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Competitors, Complementors, Parents and Places: Explaining Regional Agglomeration in the U.S. Auto Industry

Luís Cabral
New York University and *CEPR*

Zhu Wang
Federal Reserve Bank of Richmond

Daniel Yi Xu
Duke University and NBER

Working Paper No. 13-04

This draft: April 2013

Abstract. Taking the early U.S. automobile industry as an example, we evaluate four competing hypotheses on regional industry agglomeration: intra-industry local externalities, inter-industry local externalities, employee spinouts, and location fixed-effects. Our findings suggest that inter-industry spillovers, particularly the development of the carriage and wagon industry, play an important role. Spinouts play a secondary role and only contribute to agglomeration at later stages of industry evolution. The presence of other firms in the same industry has a negligible (or maybe even negative) effect on agglomeration. Finally, location fixed-effects account for some agglomeration, though to a lesser extent than inter-industry spillovers and spinouts.

JEL classification: L26; L6; R1
Keywords: Local externalities; Employee spinouts; Industry agglomeration

---

Cabral: W R Berkley Term Professor of Economics, Stern School of Business, New York University; IME and PS-PS (IESE) Research Fellow; and Research Fellow, CEPR (London); luis.cabral@nyu.edu. Wang: Research Department, Federal Reserve Bank of Richmond; zhu.wang@rich.frb.org. Xu: Department of Economics, Duke University; daniel.xu@duke.edu.

We thank Ed Glaeser, Boyan Jovanovic, Bob Lucas and participants at the NBER Summer Institute, the 2012 AEA meetings, the 2012 SED meeting and various seminars for helpful comments; and Christian Hung for excellent research assistance. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1. Introduction

Why does economic activity concentrate in cities and more generally in clusters of innovation and growth? Ever since Marshall's (1890) seminal work, an intense debate has developed among economists regarding the sources of agglomeration externalities. Marshall himself pointed to the positive externalities from specialization: regions with specialized production structures tend to be more innovative in that specific industry. In particular, Marshall pointed to the importance of knowledge spillovers: each firm learns from neighboring firms in the same industry.

> Employees from different firms in an industry exchange ideas about new products and new ways to produce goods: the denser the concentration of employees in a common industry in a given location, the greater the opportunity to exchange ideas that lead to key innovations.

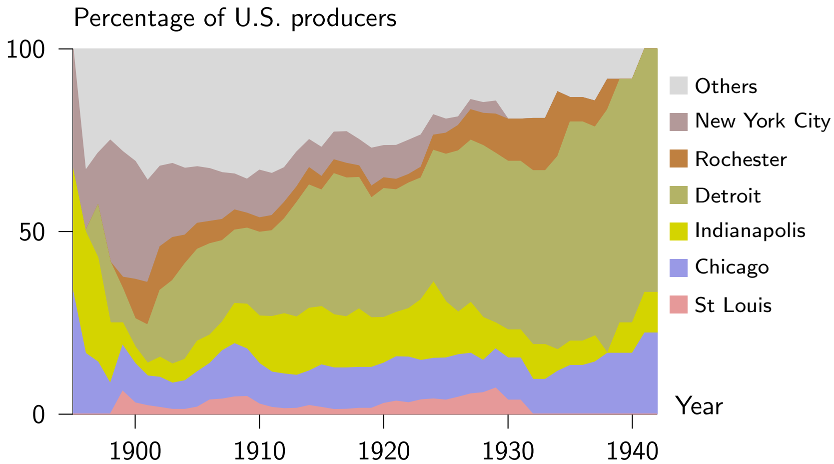
Marshall's work was later extended by the work of many authors, including Arrow (1962) and Romer (1986).[1] (Knowledge spillovers are sometimes referred to as MAR spillovers.) We should note that, in addition to knowledge spillovers, Marshall also considered other types of externalities, including input sharing and labor pooling. However, as Ellison, Glaeser and Kerr (2010) argue, all of these "predict that firms will co-locate with other firms in the same industry." Accordingly, we refer to Marshallian externalities as economic effects that lead firms to locate close to other firms of the same industry: intra-industry externalities.[2] In contrast to Marshallian externalities, other authors, most notably Jacobs (1969), proposed an alternative agglomeration thesis, the idea that knowledge spills across different industries, causing diversified production structures to be more innovative.[3] We refer to these externalities as inter-industry, or related-industry, externalities (as opposed to intra-industry, or Marshallian, externalities).

A different perspective on agglomeration is provided by the work of Klepper (2007), who focuses on the evolution of the automobile industry. He argues that "the agglomeration of the automobile industry around Detroit, Michigan is explained [by] disagreements [that] lead employees of incumbent firms to found spinoffs in the same industry."[4] The effect of spinouts on agglomeration is related to Marshallian externalities in the sense that the number of firms in a given industry is subject to self-reinforcing dynamics: the more industry i firms are located at location j, the more likely new industry i firms will be located at location j. However, the mechanism for these self-reinforcing dynamics is quite different from Marshallian externalities.

Finally, the work of Ellison and Glaeser (1997, 1999) suggests that much of the agglomeration observed in U.S. manufacturing may be due simply to the relative advantage of certain locations. For example, the wine industry is located in California, not Kansas,

1. See also Krugman (1991), Glaeser et al (1992) and Henderson et al (1995).
2. It may be argued that we are taking a very narrow view of Marshallian externalities by restricting to other firms within the same industry. In fact, we will next consider the possibility of externalities across related industries. Our goal is to present an account of factors affecting industry agglomerations that is as detailed as possible. For this reason, we believe it is helpful to make this distinction.
3. See also Jackson (1988).
4. Buenstorf and Klepper (2009) paint a similar picture for the tire industry. Note that, while Klepper (2007) uses the term "spinoffs," other authors including ourselves use the term "spinouts."

Figure 1
Geographical distribution of U.S. auto producers



largely due to California's favorable weather.[5]

In this paper, we construct a detailed dataset of the evolution of the automobile industry and run a "horse race" between alternative views of agglomeration: (a) intra-industry spillovers (Marshall et al); (b) related-industry spillovers (Jacobs et al); (c) family network, or spinout, effects (Klepper et al); (d) location fixed effects (Ellison and Glaeser).

Our analysis is motivated by Figure 1, which shows the evolution of U.S. auto production measured by each region's share of the total number of firms. The dataset that we constructed uncovers six historically important auto production centers: New York City, Chicago, Indianapolis, Detroit, Rochester, and St. Louis. This raises an important question: Why did these six locations rather than any other places become prominent auto production centers? The figure also shows that the relative importance of Detroit as a production center increased steadily since the beginning of the century. The auto industry eventually became highly concentrated in Detroit. Why?

Our results provide answers to the above questions. In contrast to Marshall's hypothesis, we find that the proximate presence of firms from the same industry has a negligible or maybe even negative effect on firm performance; we interpret this result as implying that the negative competition externality outweighs Marshallian positive spillovers. Second, confirming the ideas of Jacobs (1969) and others, we find evidence for positive spillover effects from related industries located nearby. In particular, the pre-existing carriage and wagon industry explains why the six auto production centers that we identified emerged in the first place. Thirdly, consistent with the results in Klepper (2007), we find that spinouts play a role on regional agglomeration, though mainly in a second stage of industry development. In fact, the relatively high spinout rate helps explain why the industry increasingly concentrated in Detroit over time. Finally, consistent with the work of Ellison and Glaeser (1997, 1999), we also find significant location specific effects.

Our paper's main contribution is to evaluate the relative importance of the four views outlined earlier, that is, the relative effect of competitors, complementors, parents and

5. Ellison and Glaeser (1997, 1999) also consider the possibility of "spurious" agglomeration due to non-economic motives.

places. This is possible for two reasons. First, we create a fairly complete dataset of the U.S. automobile industry that allows us to evaluate the strength of each effect by means of reduced form estimation.[6] Second, we develop a dynamic structural model that, by means of calibration and counterfactual simulation, allows us to quantify the relative contribution of each theory of industry agglomeration.[7]

In other words, our paper can be seen as an exercise in "agglomeration accounting." By analogy with the literature on growth accounting (or productivity growth accounting), we consider different explanations which are not mutually exclusive; and by looking at time series data we attempt to quantify the relative contribution of each theory to the observed rate of agglomeration. Our results, both from reduced-form regressions and from structural model calibration and simulation, suggest that intra-industry spillovers have a negligible or maybe even slightly negative effect on agglomeration; location fixed effects explain a portion of agglomeration; spinouts too have some effect, especially at later stages of industry evolution; and finally, inter-industry spillovers play a very important role. At least for the automobile industry, inter-industry spillovers seem to be the overall winning "horse."

Quantitatively, we propose the following agglomeration accounting exercise: we first calibrate our structural model to fit the data (the fit is very good). Then we conduct a series of counterfactual simulations where a selected feature of the model (e.g., spinouts) is shut off. We then take the drop in goodness of model fit as that feature's (e.g., spinouts) contribution to the overall fit. Finally, we compute the fraction that each feature represents of the total contributions of "goodness of fit" implied by each feature. The values we obtain are: intra-industry spillovers, 50.8%; spinouts, 30.3%; location fixed effects, 18.5%. (Based on the results from reduced-form regressions, we set intra-industry effects to zero in our structural model.)

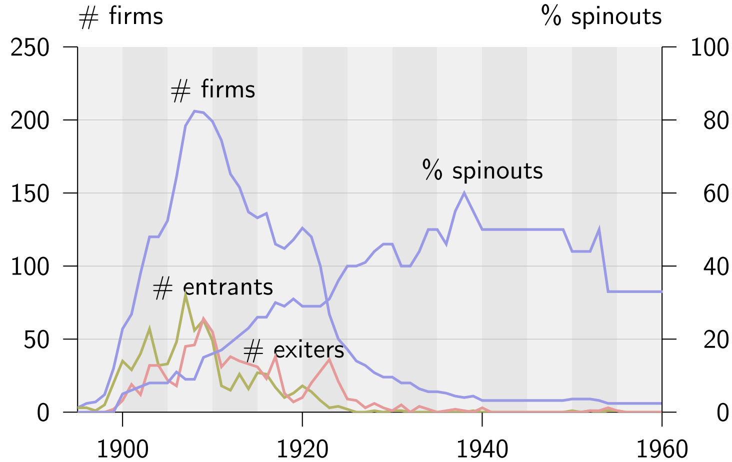
In addition to estimating the relative importance of spinouts, our results also shed light on when and how they matter: we (a) show that spinouts are only relatively important at a fairly late stage of industry evolution; (b) show that, within the spinout channel, Detroit plays a particularly important role; and (c) disentangle the "genetic" and "environmental" components of the family network effect. In other words, we revisit the old nature vs. nurture debate: is generational correlation of performance the result of good genes or of a good upbringing? Our results suggest that the positive superior performance of spinouts from high-performing parents is only observed when the spinout locates close to the parent. This suggests that the hereditary benefit may result from access to information or other forms of support rather than common "genes."

The rest of the paper is structured as follows. In Section 2, we briefly describe the evolution of the U.S. automobile industry. Next, in Section 3 we run a series of reduced-form regressions that test the relative merit of various theories of industry agglomeration. The results from these regressions motivate Section 4, where we develop and calibrate a dy-

6. Other authors, most notably Klepper (2007), have also looked at the evolution of the U.S. auto industry. In addition to data on entries, exits and spinouts, we also obtained data on the related carriage and wagon industry, which turns out to have a significant explanatory power. (To the best of our knowledge, ours is the first paper to work with data of this nature.) Moreover, unlike Klepper (2007), who only examines the evolution of Detroit as a production center, we identify six production centers and look at the entire geographic distribution of the U.S. auto industry.

7. In contrast to Klepper (2007) and others, one of our main results is that the relative contribution of each theory of agglomeration varies greatly along the industry's life cycle.

Figure 2
Evolution of the U.S. automobile industry, 1895–1969.



namic structural model of industry evolution and run a series of counterfactual simulations that allow us to quantify the relative contribution of each agglomeration factor. Section 5 concludes the paper.

2. The U.S. automobile industry

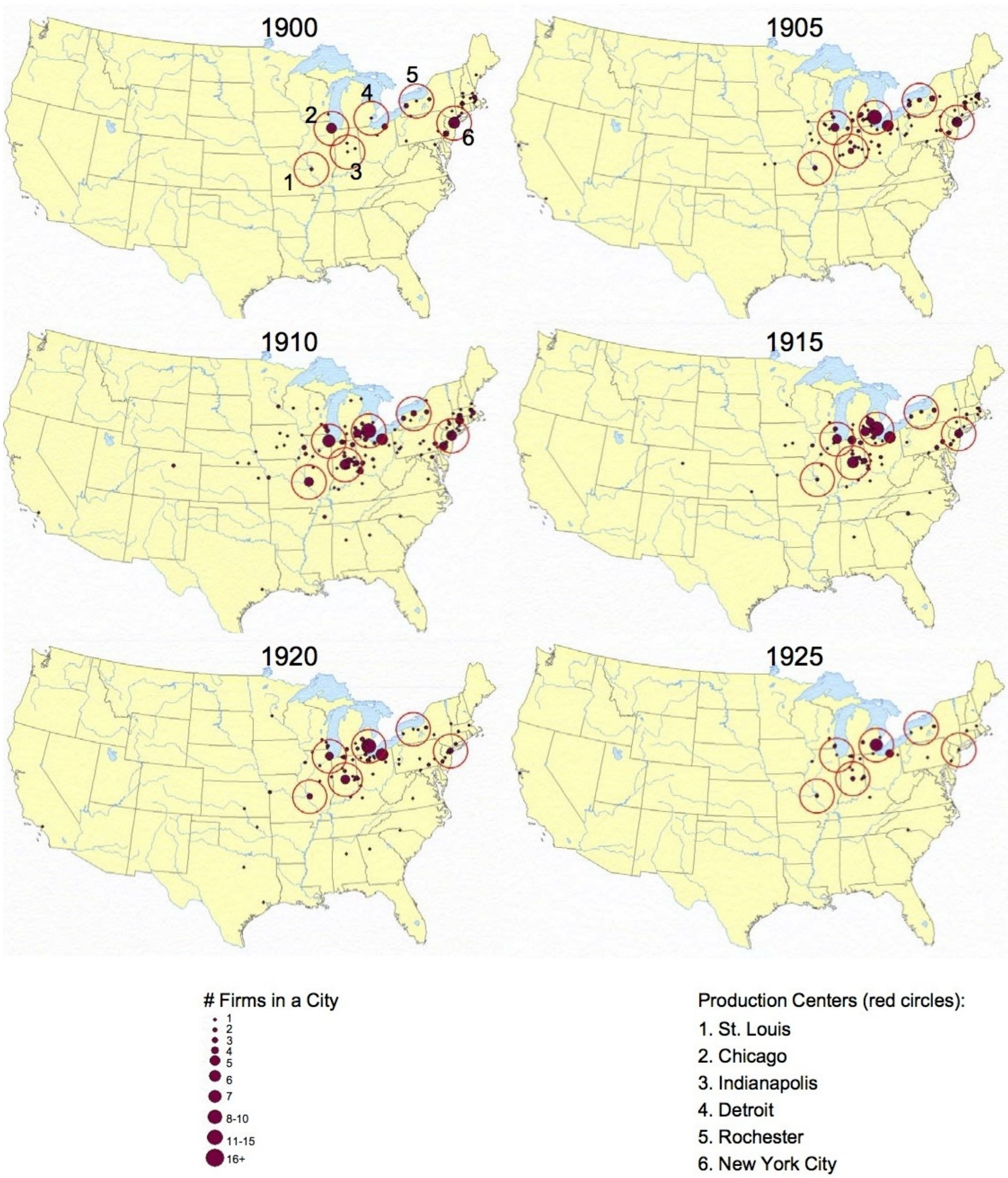
The U.S. auto industry went through tremendous development in its first 75 years, evolving from a small infant industry to a gigantic mature one. During the process, the number of firms initially rose and later fell: in its peak years around 1910, there were more than 200 producers, but only 8 survived into the 1940s. Figure 2 documents this evolution, plotting the number of active firms each year, as well as the number of entrants and exiters. As can be seen, two industry "shakeouts" took place, one around 1910 and a second one around 1920.

Figure 2 also plots the percentage of entrants due to a spinout, that is, firms that were founded by a former manager or employee of an existing automobile firm. This percentage increased from almost zero in 1900 to about 60 percent in 1940.

The automobile industry also went through substantial changes in geographic concentration patterns over the years. Figure 3 shows the number of firms in each U.S. city over a period of 25 years (from 1900 to 1925). Based on this picture, we identified six historically important automobile production centers: St. Louis, Chicago, Indianapolis, Detroit, Rochester, and New York City.[8] As shown in Figure 1, New York City and Chicago were the most important centers in the late 1890s. Soon after, Detroit and other centers caught up. By 1905, 25 percent of all active firms were located in Detroit and they produced more than 50 percent of total industry output. Meanwhile, 16 percent of the firms were located in

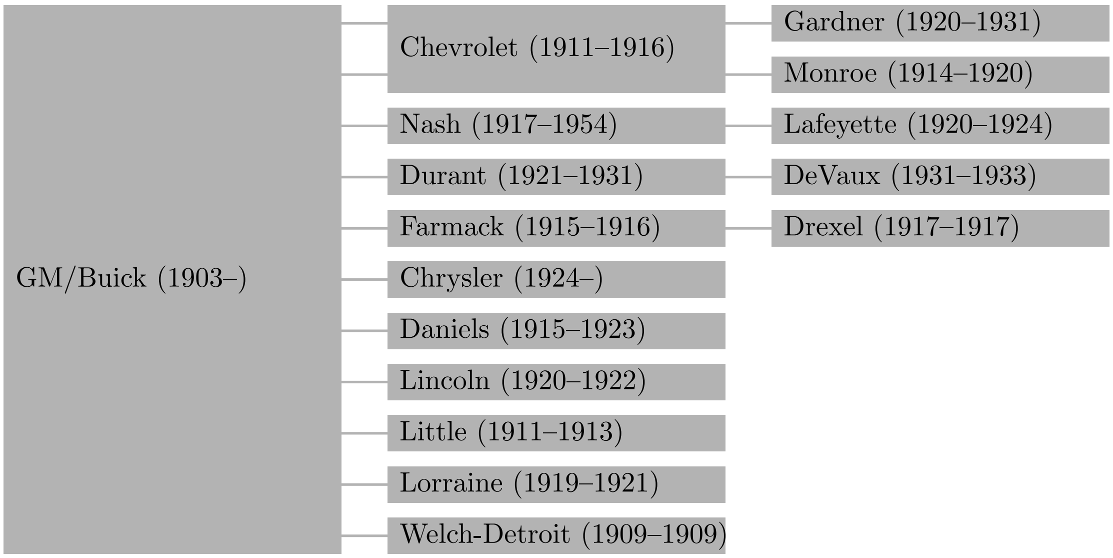
8. Lacking good data on firm size, we instead use the number of firms as a measure of regional agglomeration. A city is counted as an automobile production center city if it had at least five automobile producers in 1910 (the peak year of the auto industry in terms of firm numbers). We then define the region within 100 miles of the center city as the production center, named after the center city (we tried different radiuses ranging from 25 miles to 150 miles for the center definition, and the 100 mile radius appears to provide the overall best fit for the data).

Figure 3
Geographic concentration of U.S. auto producers (1900-1925)



New York City, 10 percent in Chicago, 8 percent in Indianapolis, 7 percent in Rochester, 2 percent in St. Louis, with the remaining 32 percent scattered across the country. Over time, Detroit gained an increasing share, both in terms of the number of firms and in terms of industry output. By 1920, 35 percent of all active firms were located in Detroit, producing

Figure 4
GM/Buick's family tree



about 70 percent of total industry output.

As mentioned earlier, an important fraction of the industry entrants originated in other existing industry firms: a spinout. We will refer to the firm originating the spinout as the "parent" and the spinout firm as a "child." Sometimes, the "child" itself becomes a parent by originating a spinout. Together, spinouts give rise to "families" of automobile firms, that is, groups of firms linked together by spinout relationships.

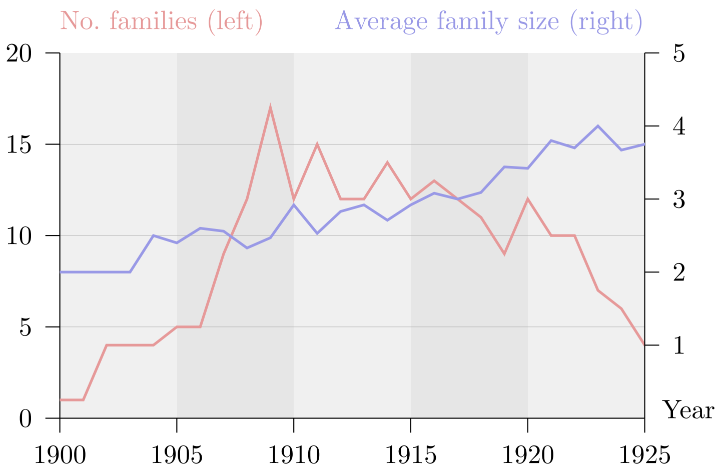
We identified a total of 53 spinout families over the history of the automobile industry. The three largest families were GM/Buick, Ford and Oldsmobile, all located in Detroit, each generating 12–17 spinouts.[9] As an example, Figure 4 displays the GM/Buick family tree — one of the largest families in the industry's history. As can be seen, it's a family with three generations: for example, a former GM employee founded Chevrolet, from which in turn Gardner and Monroe spun out.

Figure 5 plots the evolution of the number of spinout families and of family size. Early on, there were very few spinouts. For example, in 1900 only one spinout family existed (which had two members including the parent), out of a total of 57 firms in the industry. In the following two decades, the period of greatest industry turbulence (that is, higher entry and exit rates), the number of families was somewhere between 10 and 15, whereas average firm size was somewhere between 3 and 4. By 1920, 41 out of a total of 136 firms belonged to spinout families. Most spinouts located near their parents. For example, 76 percent of the spinouts in the top three families stayed in Detroit.

Although our analysis focuses on the automobile industry, there are related industries which play an important role in explaining entry and exit patterns by automobile firms. Prominent among these is the carriage and wagon industry (C&W). Figure 6 plots the level of activity in the automobile industry (measured by the number of firms in 1910, the peak

9. Klepper (2007) constructed family trees for GM/Buick, Ford, Oldsmobile and Cadillac. The family members that he identified are largely consistent with ours.

Figure 5
Family size distribution of U.S. auto producers (1900–1925)



year of the auto industry in terms of firm numbers) against the C&W industry (measured by C&W employment level in 1904).[10] As can be seen, there is a clear positive correlation between the two. Obviously, at this stage there is little more to be said than the fact there is a correlation. Below we explore this relation in greater detail, and the results suggest that location patterns of automobile firms (the new industry) may indeed have been influenced by location patterns of C&W firms (the older industry).

Take for example the case of Michigan. The C&W industry predates the automobile industry by a few decades. In 1868–1869, two Canadians each started their shop in Flint, Michigan. Many others followed, and before long there were more than 100 C&W firms located in various centers across the state, including Detroit.[11] In 1898, the 50 largest Michigan companies produced more than 300,000 horse-drawn vehicles.

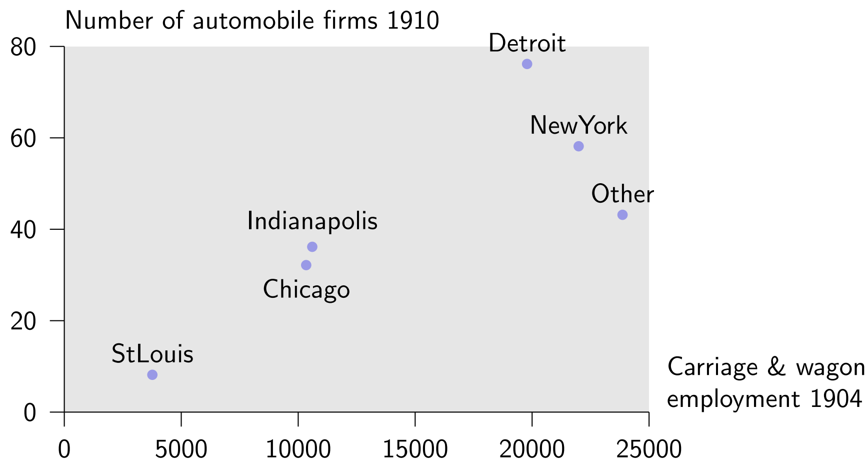
Although it is not our present focus, it is worth mentioning that one important factor leading to the concentration of C&W firms in Michigan was the presence of an important input: wood. Obviously, there were other locations in the U.S. with equal or better access to this input, and indeed the C&W industry was present in other locations as well.

Why is the C&W industry important in studying the automobile industry? Anecdotal evidence suggests that many automobile firms were founded by experienced C&W veterans. For example,

> In 1895 William C. Durant (a grandson of Henry Crapo, a leading figure in the lumbering era and governor of Michigan from 1865 to 1869) along with J. Dallas Dort and W. A. Patterson incorporated their business as the Durant-Dort Carriage Company. Through creating and developing this business, Durant learned skills that he employed later in the automobile industry.[12]

10. Given that only state-level data are available for the C&W industry employment, we combine New York City and Rochester together as New York in the figure. We also add up 30 non-center states in the "other" group.
11. About three-quarters of the C&W workers were concentrated in Kalamazoo, Jackson, Grand Rapids, Detroit, Pontiac, Lansing, and Flint.
12. See `http://www.michigan.gov/dnr/0,4570,7-153-54463_18670_18793-68566--,00.html`, accessed May 30, 2012.

Figure 6
Carriage and wagon industry



3. Reduced-form regression analysis

In this section we present a series of reduced-form regressions that provide tests of various theories of industry agglomeration. Marshallian theories imply that a firm's benefit from locating in region i is increasing in the number of other firms located in region i. Everything else constant, we would expect this to be reflected in entry rates and exit rates: entry rates are increasing in the numberer of firms, whereas exit rates are decreasing in the number of firms in region i.

Regarding co-agglomeration economies, our tests are based on data regarding the importance of the carriage and wagon industry, an industry related to the automobile industry. The theory prediction is that the presence of related industries improves a firm's prospect. We thus expect entry (respectively, exit) rates to be increasing (respectively, decreasing) in the presence of the carriage and wagon industry. As a very preliminary result, Figure 6 plots the number of automobile firms against employment levels in the carriage and wagon industry. The scatter plot suggests a clear positive correlation between the two variables. Below we present more systematic evidence on the relation between these variables as a test of the co-agglomeration hypothesis.

As shown by Klepper and others, spinouts are an important factor in the development of a new industry, especially at a second stage of industry evolution (by definition, the first entrant cannot be a spinout). Spinouts per se do not imply agglomeration: if every incumbent firm is equally likely to generate a spinout, then the fraction of industry firms accounted for by region i does not change as a result of spinouts. The question is then whether spinout rates vary systematically from region to region or as a function of region size.

■ **Data.** Our data comes from various sources. First, Smith (1970) provides a list of every make of automobile produced commercially in the United States from 1895 through 1969. 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

Table 1
Firm level summary stats

Variable	Obs	Mean	St Dev	Min	Max
De novo entrant	771	0.30	0.46	0	1
De alio entrant	771	0.52	0.50	0	1
Spinout entrant	771	0.17	0.38	0	1
Top firm	771	0.06	0.24	0	1
Entry year	771	1908	6	1895	1939

underwent. Smith's list of car makes was used to derive entry, exit and location of firms.[13]

Second, Kimes (1996) provides comprehensive information for every auto make produced in the U.S. from 1890 through 1942. Using Kimes (1996), we were able to collect additional biographical information about the entrepreneurs who founded and ran each individual firm. An entrepreneur was categorized into one of the following three groups: de novo, de alio, or spinout entrants. De alio entrants are firms whose founder had prior experience in related industries before starting an automobile firm. Spinouts are firms whose founders previously worked as employees in existing automobile firms. Finally, de novo entrants includes all other entrants, those firms whose founders had no experience in either autos or related industries. Kimes' information was used to derive family linkages between individual firms. In other words, we constructed family trees for spinout firms.

The third data source is Bailey (1971), which provides a list of leading automobile makes from 1896–1970 based on annual sales, specifically, the list of top-15 makes. We used this information to identify top automobile producers during these periods.

Additionally, we collected information on regional economy as well as auto-related industries across U.S. states, including per capita income, population, carriage and wagon industry employment, and production. These data come from various issues of the *Census of the U.S. Manufactures* and the *Statistical Abstract of the United States*.

■ **Regression variables and descriptive statistics.** Our dataset includes every U.S. company that ever sold at least one passenger car to the public during the first 75 years of the industry (1895–1969), a total of 775 firms. In our following reduced-form regression analysis, we set our sample range up to WWII (1895–1942), which includes 771 firms.[14]

Tables 1 through 3 provide summary statistics for the variables used in the regressions. Table 1 includes firm-level variables, as follows:

- De novo entrant. Equals 1 if the firm's founder had no experience in the automobile or related industries. 30% of all entrants were de novo entrants.
- De alio entrant. Equals 1 if the firm's founder had previous experience in an automobile-related industry. 52% of all entrants were de alio entrants.

13. The entry and exit are based on the first and last year of commercial production.
14. Klepper (2007) used similar data sources to study the concentration of U.S. auto production in Detroit. Our data is collected independently but comparable to Klepper's in many aspects. However, we take a different approach and study broader questions.

Table 2
Firm-year level summary stats

Variable	Obs	Mean	St Dev	Min	Max
Firm age	4454	6.87	7.24	1	43
Firm death	4454	0.17	0.38	0	1
Spinout birth	4454	0.02	0.13	0	1
Family size	4454	1.53	1.52	1	10
Family top	4454	.47	1.04	0	5
Local family size	4454	1.37	1.25	1	9
Local family top	4454	.41	.94	0	5
Non-local family size	4454	.16	.70	0	9
Non-local family top	4454	.07	.39	0	5
Center size	4454	35.79	23.43	1	96
Center top	4454	6.03	5.55	0	18

- Spinout entrant. Equals 1 if the firm's founder had previous experience in the automobile industry. 17% of all entrants were spinout entrants.
- Top firm. Equals 1 if, at any point during its life, a firm was one of the top automobile producers, as classified by Bailey (1971). Only 6% of the 771 firms are in this category.
- Entry year. First year when the firm started commercial production. Varies from 1895 to 1939, with an average of 1908.

We next turn to firm-year level variables. We define location dummies corresponding to St. Louis, Chicago, Indianapolis, Detroit, Rochester, New York City and others. The summary statistics of firm-year level variables are listed in Table 2.

- Firm age. Difference between current year and entry year. It ranges from 1 to 43 in our sample. The average is 6.87 years, not very different from what is found in other industries.
- Firm death. Equals 1 if the firm stops commercial production during the current year. The average 0.17 corresponds to a hazard rate somewhat higher than that found in other industries, but one must remember that we are looking at the initial stages of a new industry, where entry and exit rates are typically higher.
- Spinout birth. Equals 1 if a firm generates a spinout entrant in current period, that is, a firm employee founds a new firm in the automobile industry. The average spinout birth hazard rate is about 2%.
- Family size. Number of firms belonging to the firm's family (including itself) in the current period. On average, a firm belongs to a family of 1.53 firms; the minimum is 1 and the maximum 10.

Table 3
Region level summary stats

Variable	Obs	Mean	St Dev	Min	Max
Population 1900	35	2126.14	2895.87	245	16390
Per capital income 1900	35	187.4	87.83	72	415
C&W employment 1904	35	2579.77	5188.21	23	21991

- Family top. Number of top firms belonging to the firm's family (including itself) in the current period. On average, there are .47 top firms in a firm's family; the minimum is zero and the maximum 5.
- Local family size. Number of firms belonging to the firm's family (including itself) in the current period that are located in the same region as the firm in question. The average is 1.37, a little lower than 1.53 firms, suggesting that a firm typically locates close to its family.
- Local family top. Number of top firms belonging to the firm's family (including itself) in the current period that are located in the same region as the firm in question.
- Non-local family size. Number of firms belonging to the firm's family (including itself) in the current period that are not located in the same region as the firm in question.
- Non-local family top. Number of top firms belonging to the firm's family (including itself) in the current period that are not located in the same region as the firm in question.
- Center size. Number of firms in a given location and year. It varies from 1 to 96 and has an average of about 36 firms.
- Center top. Number of top firms in a given location and year. It varies from 0 to 18 and has an average of about 6 firms.

Finally, we have the following region-level variables. Given that only state-level data are available, we regroup the production centers and define 35 regions: St. Louis, Chicago, Indianapolis, Detroit, New York (combining New York City and Rochester), and 30 other states. The summary statistics are listed in Table 3:

- Population 1900. Regional population (thousands) in 1900.
- Per capita income 1900. Regional per capita income in 1900.
- C&W employment 1904. Number of workers in the C&W industry in a given region in 1904.

■ **Marshall vs Jacobs: entry by non-spinout firms.** Table 4 presents our first set of regressions where the dependent variable is the number of non-spinout entrants in location i. The sample range is 1895-1915 (the period when most entry takes place) and we consider 35 different locations: 5 production centers and 30 other states, for a total 735 location-year observations.[15] All regressions are based on a negative binomial model. The different

15. Due to data limitation, we combine New York City and Rochester into one center (New York).

Table 4
Negative binomial models of non-spinout entry.
Dependent variable: number of non-spinout entrants in location i

	Spec 1	Spec 2	Spec 3	Spec 4	DeNovo	DeAlio
Log polulation 1900	1.028*** (0.18)	0.172 (0.26)	0.297 (0.34)	0.251 (0.33)	-0.022 (0.33)	0.178 (0.50)
Log per capita income 1900	1.236*** (0.34)	1.150*** (0.26)	1.767*** (0.35)	1.724*** (0.34)	1.758*** (0.47)	1.824*** (0.42)
Center size (t-1)	0.017*** (0.00)	0.015*** (0.00)	-0.004 (0.01)		0.004 (0.01)	-0.006 (0.01)
Log C&W employment 1904		0.671*** (0.17)	0.896*** (0.21)	0.889*** (0.20)	1.022*** (0.23)	0.991*** (0.29)
Year	-0.012 (0.01)	-0.006 (0.01)	0.058** (0.03)	0.045*** (0.02)	0.031 (0.04)	0.064** (0.03)
Constant	7.937 (26.87)	-0.469 (27.05)	-127.949** (51.37)	-102.28*** (32.52)	-75.608 (78.02)	-138.98** (59.68)
Sample	1895–1915	1895–1915	1900–1910	1900–1910	1900–1910	1900–1910
Log Likelihood	-475	-469	-284	-284	-174	-231
N	700	700	385	385	385	385

Notes: Standard errors in parentheses. Star levels: 10, 5 and 1%.

models refer to different sets of independent variables (specifications 1 to 4) and different dependent variables (columns 5 and 6).

The first regression suggests that population, per capita income and center size are all positively correlated with the entry rate of non-spinout firms in a given region. In particular, the third coefficient is constant with the prediction of Marshall's model of agglomeration economies. The second regression adds the regional level of employment in the carriage and wagon industry as an explanatory variable. In this second model, population in region i is no longer statistically significant. The center size coefficient is still significant and has a very similar value.

Our next two regressions restrict the sample to the 1900–1910 period. The main reason for this alternative sample specification is that, due to data limitations, we can only use 1904 data on C&W employment levels. As such, our measurement error is likely to be high as we consider dates far from 1904. Under this alternative specification, we observe that the coefficient on C&W employment is still significant but that of center size is not (specification 3). Excluding center size does not change the coefficient of C&W employment that much. More important, excluding center size does not affect the model fit. In fact, the log likelihood remains the same, which further suggests that center size is not an important entry determinant.

Our last two models split the dependent variable into de novo and de alio entry. The results suggest that for both types of entry both per capita income and C&W employment are associated with higher entry rates. The center size coefficient, however, is not significant.

In the above set of regressions, the "horse race" between Marshall-type agglomeration economies and Jacobs-type co-agglomeration economies is proxied by the "race" between center size and C&W employment as explanatory variables. The discussion above suggests that the statistical power of center size is weaker than that of C&W employment as an explanatory variable. We next argue that the economic significance is also much lower. One way of measuring the relative importance of independent variable x is to compute the value of $\beta_x \sigma_x$, that is, the product of the regression coefficient and the variable's standard deviation. Taking the specification 2 as an example, this comes to $0.015 \times 12.05 = .18$ for the auto industry center size variable and $0.671 \times 1.63 = 1.09$ for the C&W employment level variable: almost one order of magnitude higher.

All in all, our first set of regressions provides little support for Marshall-type economies but considerable support for the idea of related industries as a source of co-agglomeration economies.

■ **Are spinouts an agglomeration force?** About 17% of all firm entries in the history of the U.S. auto industry correspond to managers or employees of existing firms who leave the company to start their own: a spinout. To the extent that spinout rates vary across regions, it is conceivable that spinouts may act as a force toward agglomeration. We now consider a series of regressions to test this possibility.

Table 5 presents the results of six model specifications, which differ in the set of independent variables considered. The level of observation is firm-year and the sample range 1895–1942, which results in a total of 3,000 to 4,000 observations approximately (depending on the set of independent variables included). Some of the independent variables — center size, family size, center top and family top — are lagged one year. In this way, we avoid including the new entrants in the measure of existing firms.[16]

Some broad patterns emerge from this set of regressions. First, firm age has a positive and significant coefficient throughout. This suggests that older firms are more likely to give birth to a spinout than younger firms. Note that we do not include direct measures of firm quality on the right-hand side. For this reason, we expect firm age to capture firm ability to some extent, in which case the results suggest that higher ability firms are more likely to give birth to a spinout.

In addition to the "parent's" characteristics (for which age is a proxy), the likelihood of giving birth to a spinout is also a function of the parent's family. Specifications 3 and 5 suggest that the greater a firm's family size, the more likely the firm will give birth to a new spinout. In specifications 4 and 6 we use family top instead of family size. This alternative specification places extra weight on the quality of parent's family. The coefficient remains significant. It is higher in value, though we should add that both the average and standard deviation of family top is lower than those of family size.

Similarly to Table 4, center size does not seem to have a significant result. Center top, which controls for the quality of the firms in a given geographic location, is significant in one specification but the coefficient is negative and small.

In the first four specifications, we use year as an explanatory variable. The positive coefficient reflects the fact that spinouts are more likely during later stages of an industry evolution (by definition, the probability of a spinout is zero until the first firm enters). Since the evolution is not linear, in specifications 5 and 6 we use year dummies instead of the variable year.

Last but not least, the various center dummies seem not to be significant except for

16. The regression results are very similar if contemporary values are used.

Table 5
Logit models of spinout entry.
Dependent variable: firm gives birth to spinout

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6
Firm age	0.062*** (0.02)	0.065** (0.03)	0.072*** (0.02)	0.056** (0.02)	0.090*** (0.03)	0.079*** (0.03)
Center size		0.015 (0.01)	0.012 (0.01)		-0.020 (0.01)	
Family size			0.168*** (0.05)		0.174*** (0.05)	
Center top				0.042 (0.08)		-0.183** (0.09)
Family top				0.291*** (0.09)		0.315*** (0.09)
Chicago	0.608 (0.47)	1.308** (0.66)	1.180* (0.64)	0.797 (0.51)	-0.018 (0.60)	0.256 (0.49)
Indianapolis	0.579 (0.46)	1.027 (0.63)	0.926 (0.61)	0.618 (0.50)	-0.300 (0.60)	-0.064 (0.53)
Detroit	1.466*** (0.40)	1.744*** (0.46)	1.447*** (0.45)	0.797 (0.78)	1.110*** (0.40)	3.063*** (0.96)
Rochester	0.255 (0.58)	1.067 (0.80)	0.911 (0.77)	0.489 (0.60)	-0.505 (0.74)	-0.112 (0.58)
New York City	0.738 (0.48)	1.425** (0.66)	1.250** (0.63)	0.800 (0.49)	0.242 (0.57)	0.743 (0.49)
Constant	104.562*** (40.77)	117.157*** (45.35)	146.479*** (46.76)	143.176*** (44.70)	-3.838*** (1.16)	-4.472*** (1.03)
Year or year dummies	-0.057*** (0.02)	-0.064*** (0.02)	-0.080*** (0.02)	-0.078*** (0.02)	D	D
N	4333	3585	3585	3585	2979	2979

Notes: Center Size, Family Size, Center Top and Family Top one-year lagged.
Robust standard errors clustered by firm. Star levels: 10, 5 and 1%.

Detroit.[17] In 5 out of our 6 specification the coefficient is statistically significant. Moreover, the coefficient's size is also quite significant in comparison to other determinants of spinout birth. For example, being in Detroit increases the probability of giving birth to a spinout by more than a 5 standard deviation increase in firm age or adding 9 top firms to a firm's family (specification 6). In terms of odds ratio, being in Detroit increases the odds ratio of giving birth to a spinout by more than 21 times (specification 6).

As robustness checks, we also ran random-effects logit models and considered different sample ranges (e.g., 1895–1929). The results are all similar.

■ **Survival of the fittest: determinants of exit rates.** As happens in many industries, net entry and exit rates in the automobile industry are considerably lower than gross entry and exit rates. Consequently, understanding exit patterns is an important step towards understanding the evolution of industry concentration. Our next set of results pertains

17. St. Louis is omitted in the regressions due to no spinout.

Table 6
Logit models of firm exit. Dependent variable: firm exit

	Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6
Firm age	-0.051*** (0.01)	-0.070*** (0.01)	-0.068*** (0.01)	-0.057*** (0.01)	-0.073*** (0.01)	-0.063*** (0.01)
Center size		0.003 (0.00)	0.003 (0.00)		0.001 (0.00)	
De Alio			-0.467*** (0.11)	-0.462*** (0.10)	-0.500*** (0.11)	-0.492*** (0.11)
Spinout			-0.452*** (0.15)	-0.214 (0.16)	-0.451*** (0.16)	-0.226 (0.16)
Family size			-0.081** (0.04)		-0.098** (0.04)	
Center top				0.026 (0.03)		0.029 (0.03)
Family top				-0.373*** (0.08)		-0.396*** (0.08)
St Louis	-0.024 (0.26)	0.163 (0.33)	0.250 (0.32)	0.209 (0.32)	0.013 (0.34)	0.113 (0.31)
Chicago	-0.143 (0.15)	-0.062 (0.19)	-0.074 (0.18)	-0.146 (0.17)	-0.186 (0.22)	-0.155 (0.17)
Indianapolis	-0.532*** (0.13)	-0.457*** (0.16)	-0.385** (0.16)	-0.504*** (0.16)	-0.497*** (0.18)	-0.522*** (0.17)
Detroit	-0.370*** (0.11)	-0.395*** (0.12)	-0.271** (0.12)	-0.442 (0.29)	-0.298** (0.12)	-0.476 (0.34)
Rochester	-0.177 (0.21)	0.064 (0.25)	-0.055 (0.27)	-0.152 (0.25)	-0.211 (0.28)	-0.190 (0.24)
New York City	0.243** (0.11)	0.316** (0.16)	0.320* (0.17)	0.232* (0.14)	0.232 (0.19)	0.236* (0.14)
Constant	-50.496*** (12.32)	-52.794*** (14.58)	-62.052*** (15.14)	-55.111*** (14.40)	-0.738 (0.49)	-0.823* (0.47)
Year or year dummies	0.026*** (0.01)	0.027*** (0.01)	0.032*** (0.01)	0.028*** (0.01)	D	D
N	4454	3683	3683	3683	3602	3602

Notes: Center Size, Family Size, Center Top and Family Top one-year lagged.
Robust standard errors clustered by firm. Star levels: 10, 5 and 1%.

precisely to firm exit.

Table 6 displays the results of six different logit regressions. In all of them, the dependent variable is firm exit, that is, a dummy variable that takes the value 1 if a firm exits in a given year. Different specifications correspond to different sets of independent variables. Some of the independent variables — center size, family size, center top and family top — are lagged one year.[18]

In all regressions, firm age has a negative coefficient. This is consistent with much of the previous literature on firm exit: older firms are less likely to exit than younger firms.

18. The regression results are very similar if contemporary values are used.

Specifically, a one-standard deviation increase in firm age decreases the odds ratio by about 41% (specification 5).

Other things being equal, de alio and spinout firms are less likely to exit than de novo firms. Specifically, the odds ratios of exit are about 39% lower for de alio firms and 36% lower for spinout firms with respect to de novo firms (specification 5).

We saw earlier that center size does not seem to have a big impact on firm entry (cf Tables 4 and 5), Table 6 suggests that it does not have an impact on firm exit either. Again, the evidence does not seem to match the prediction of Marshall-type agglomeration economies.

By contrast, family size and family top seem to have a significant impact on survival: an additional "relative" (that is, an additional firm in a firm's family) reduces the odds ratio of exit by about 9% (specification 5), whereas an additional top relative reduces the odds ratio of exit by about 33% (specification 6).

Finally, the various center dummies suggest that there are significant location specific effects, with firms in Indianapolis and Detroit more likely to survive than firms in other production centers.

■ **All in the family: determinants of spinout peformance.** Several of the above regressions suggest that "family matters." Specifically, the size and quality of a family has an important impact on whether a spinout will take place and whether such spinout will survive. We now take a closer look at the mechanism whereby family membership helps the survival of a spinout firm. As in many other settings, an interesting question is the split between "nature" and "nurture": do spinouts perform better because they are helped by their parents (nurture) or simply because they have better genes (nature)?

Table 7 displays four logit regressions where the dependent variable, as in the previous table, is firm exit, that is, a dummy variable that takes the value 1 if a firm exits in a given year. Different specifications correspond to different sets of independent variables. Some of the independent variables — center size, family size, center top and family top — again are lagged one year.[19] Differently from the regressions in Table 6, we now split the family size and family top variables by location: local family size now measures the number of relatives in the same location, whereas non-local family size measures the number of relatives located elsewhere (a similar distinction applies to local and non-local family top).

The results are quite striking: whereas the local variables (family size and family top) are statistically significant, the non-local ones are not statistically significant. In terms of coefficient size, local family top shows greater values (in absolute terms) than family top in Table 6. In fact, when we include local family top (specifications 2 and 4 in Table 7) the variable spinout ceases to be statistically significant. This suggests that belonging to a family of high performance firms and being located nearby family relatives is associated with superior spinout performance, whereas if a firm is located far from its family then performance is not statistically different from that of de novo entrants.[20] In other words, the results suggest that, in the case of the automobile industry, nurture trumps nature.

19. The regression results are very similar if contemporary values are used.

20. To investigate the potential endogeneity of a spinout's location choice, we collected additional information on the motive of each spinout from the top three families: GM/Buick, Ford and Oldsmobile. The results indicate no systematic bias between the motive of a spinout and its location choice.

Table 7
Logit models of firm exit. Dependent variable: firm exit

	Spec 1	Spec 2	Spec 3	Spec 4
Firm age	-0.068*** (0.01)	-0.053*** (0.01)	-0.073*** (0.01)	-0.059*** (0.01)
De Alio	-0.467*** (0.11)	-0.463*** (0.10)	-0.500*** (0.11)	-0.492*** (0.11)
Spinout	-0.455*** (0.15)	-0.230 (0.16)	-0.451*** (0.16)	-0.237 (0.16)
Center size	0.003 (0.00)		0.001 (0.00)	
Local family size	-0.091** (0.04)		-0.100** (0.05)	
Non-local family size	-0.060 (0.08)		-0.094 (0.08)	
Center top		0.028 (0.03)		0.031 (0.03)
Local family top		-0.465*** (0.10)		-0.486*** (0.11)
Non-local family top		-0.111 (0.13)		-0.147 (0.15)
St Louis	0.243 (0.32)	0.135 (0.32)	0.011 (0.34)	0.035 (0.32)
Chicago	-0.075 (0.18)	-0.148 (0.17)	-0.187 (0.21)	-0.161 (0.17)
Indianapolis	-0.381** (0.16)	-0.488*** (0.16)	-0.496*** (0.19)	-0.511*** (0.17)
Detroit	-0.262** (0.12)	-0.411 (0.29)	-0.296** (0.13)	-0.439 (0.34)
Rochester	-0.053 (0.27)	-0.144 (0.25)	-0.211 (0.28)	-0.183 (0.24)
New York City	0.321* (0.17)	0.238* (0.14)	0.233 (0.19)	0.239* (0.14)
Constant	-61.233*** (15.16)	-48.913*** (14.47)	-0.736 (0.49)	-0.819* (0.47)
Year or year dummies	0.032*** (0.01)	0.025*** (0.01)	D	D
N	3683	3683	3602	3602

Notes: Center Size, Family Size, Center Top and Family Top one-year lagged.
Robust standard errors clustered by firm. Star levels: 10, 5 and 1%.

■ **Robustness analysis and further notes.** We performed a series of robustness checks on our results of firm exit. First, we considered alternative treatments of exit. In our exit regressions, we did not separate exit by acquisition from exit by liquidation. It may be argued that exit by being acquired should not be counted as firm failure (in fact, it may be quite the opposite). In our sample, we have 762 exits in total, of which 108 (14%) resulted from acquisition. One way to solve the potential problem of confounding the two types of exit is to count exits by acquisition as censored observations of exits. When we run these alternative regressions, the results are similar in value and stronger in statistical significance.

Second, in our spinout and exit regressions we did not include top firm as an explanatory variable. The reason is that we would like the explanatory variables (e.g. center size, family size, etc) to predict the firm's performance in terms of spinout and exit. Top firm is just another measure of firm performance, which duplicates the dependent variables. Of course, top firm itself could be an imperfect proxy for firm performance (the way we define it as ever being a top firm). Therefore, when we do include it as an explanatory variable, our results still hold in terms of coefficient signs and values but become statistically weaker.

Third, we did not include C&W employment in our spinout and exit regressions. One reason is that we only have C&W data for 1904.[21] Another reason is that we can use family size or family top to proxy for firm quality (as our following theory explains). As such, the effect of C&W employment on firm quality is indirectly controlled.

Finally, we ran random-effects logit models and considered different sample ranges (e.g., 1895–1929). The results are all similar.

■ **Summary of main empirical results.** We may summarize our empirical findings as follows:

- Entry rates by non-spinout firms (that is, de novo and de alio entrants) in a given location are increasing in: (a) regional income; (b) regional employment levels in the C&W industry.
- Spinout entrants are more likely to come out of: (a) older firms; (b) larger and better families; (c) Detroit.
- Firm survival rates are higher if: (a) the firm is older; (b) the firm resulted from de alio entry; (c) the firm was spun out of a high-performance parent and remained in the same location as the parent; (d) the firm is located in Detroit or Indianapolis.
- Firm entry and survival rates do *not* depend on center size. If anything, entry rates are negatively impacted by center top (cf Table 5, specification 6).

Taken together, the evidence casts doubt on the importance of Marshall (1890) type agglomeration economies. By contrast, it suggests that Jacobs (1969) type co-agglomeration economies may play an important role, as well as Klepper (2005) type spinout entry (especially at a later stage of industry development). Finally, as suggested by the work of Ellison and Glaezer (2000), the results also unveil some significant location fixed effects, including, but not exclusively, Detroit.

The reduced-form analysis is useful for getting a first glance at the sign and size of the various effects. It also provides a useful springboard for our next step: to develop and calibrate a structural model of the U.S. automobile industry. Such a model will allow us to perform a series of counterfactual exercises that will better evaluate the relative weight of each force of industry agglomeration.

4. A structural model of the U.S. automobile industry

In this section, we develop a simple model of industry dynamics in the spirit of Hopenhayn (1992). Firms are forward looking, competitive price takers producing a homogeneous product with heterogeneous production capabilities. We consider two types of entrants: *de*

21. The census report is available every 5 years. However, the 1899 C&W data covers very few states, and after the auto industry took off the C&W industry declined substantially.

novo entrants and *spinout* entrants.[22] De novo entrants originate outside of the industry, that is, their founder was not an industry participant before founding the firm. Spinout entrants, by contrast, are founded by former industry participants, that is, managers or workers previously employed by an existing industry participant.

We assume that the "supply" of de novo entrants is determined by traditional channels such as local population, income and the presence of related industries. Regarding spinout entrants, we assume that each incumbent firm generates *potential* spinouts at a constant per period rate. However, just as with de novo entrants, a potential spinout makes an optimal entry decision. In other words, every actual and potential firm is treated as a rational, forward-looking agent who makes optimal entry and exit decisions.

We also make the assumption that a spinout shares the same capability with its parent firm, that is, firm capabilities are "hereditary." However, the model does not restrict the source of the hereditary effect: the family specific capability may reflect common "genes," or it could result from interactions of member firms within the family network (e.g., through knowledge linkages or business relations).

Finally, consistent with the empirical evidence presented in the previous section, we assume there are no intra-industry effects other than through family effects.

■ **Individual firm's problem.** The model is cast in discrete time and infinite horizon. A continuum of firms produce a homogenous good in a competitive market. Each firm is indexed by its discrete capability $s \in \{0, 1, ..., \bar{s}\}$ and location j. For simplicity, we assume that a firm with capability s starting at location j will retain the same capability and operate at the same location for the rest of its life. The industry structure is thus summarized by $m(s, j)$, the total mass of firms of capability s at location j. Given our assumption regarding capability and location, the evolution of $m(s, j)$ is entirely governed by entry and exit, the main focus of our analysis.

In each period, incumbent firms engage in product market competition by taking industry price p as given. Each firm chooses optimal output $q(s; j, p)$ based on its capability and location characteristics. Their period profit is denoted by $\pi(s; j, p)$. We assume $q(s; j, p)$ and $\pi(s; j, p)$ are continuous, bounded, and strictly increasing in s and p.

Once an incumbent firm obtains its profit, it decides whether to continue operating or instead to leave the industry and earn an outside options ϕ^x. The value of the outside option is privately known by the firm and i.i.d. according to cdf $F(\phi^x)$. Given its belief of a time-series sequence of industry price $\bar{p}$, an incumbent's problem can be defined as:

$$V(s; j, \bar{p}, \phi^x) = \pi(s; j, p) + \max\{\, VC, \phi^x \}$$

where the value of continuation is

$$VC(s; j, \bar{p}) = \beta \int V(s; j, \bar{p}, \tilde{\phi}^x)\; dF(\tilde{\phi}^x)$$

Potential entrants at each location make their entry decisions at the same time as incumbents. As mentioned earlier, we consider two types of entrants: de novo and spinout. De

22. For simplicity, we conflate de novo and de alio entry into one single category: de novo. We do so for two reasons. First, it keeps our structural model simpler. Second, to the extent that we use C&W to estimate the quality of non-spinout entry in each location (cf Table 8: being a top non-spinout entrant), we effectively allow for the effect of location differences to influence the average quality of non-spinout entrants.

novo entrants originate outside the industry. We assume the total mass of potential de novo entrants at location j, M_j, is determined by location specific characteristics. Each potential de novo entrant is endowed with a sunk entry cost ϕ^e. If the potential entrant pays ϕ^e then it is given an initial draw of capability s from the distribution $\mu(s,j)$, the discrete density function of capability s at location j. Hence, a potential entrant's probability of entry is given by Ψ_j, the probability that the ex-ante expected value of incumbency is greater than the entry cost ϕ^e. It follows that the expected number of de novo entrants at location j is given by

$$n_j = \Psi_j \, M_j = \Pr\left(\sum_s VC(s;j,\bar{p}) \, \mu(s,j) \geq \phi^e \right) M_j \tag{1}$$

The second type of entrants, spinouts, originate within the industry. Each period, an incumbent firm at location j has a probability γ_j of generating a potential spinout. We assume the potential spinout shares the same capability s with its parent and knows its capability when making the entry decision. As a result, the spinout's entry decision is equivalent to its parent's continuing decision: A potential spinout will enter if its entry value is higher than its random outside option ϕ^x, i.e.,

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

We assume that if a potential spinout entrant chooses not to enter in the current period, then the opportunity is foregone forever.

Note that there are two important differences between the two types of entrants. First, while potential de novo entrants are uncertain about their capability of operating in a new industry, spinout entrants directly inherit their parent's capability draw. This is a sharp assumption we make to highlight the fact that spinout entrants have better knowledge of their own capability given by their industry experience. Second, we assume that de novo entrants need to pay an additional entry cost ϕ^e with respect to spinouts, a difference that corresponds to the extra investment de novo entrants need to make to build up business relations or a customer base in the industry.

■ **Supply and demand.** We next derive the transition of the mass of firms of capability s at location j. This transition depends on the number of exits, spinouts, and de novo entrants at each state (s,j). Specifically, we have

$$m'(s,j) = m(s,j)\,(1+\gamma_j)\,\chi_{s,j} + n_j \, \mu(s,j) \tag{2}$$

where

$$\chi_{s,j} = F\big(VC(s;j,\bar{p})\big) \tag{3}$$

is the probability of staying in the industry given the cdf function F of the outside option.

The right-hand side of (2) reflects the two sources of entry mentioned earlier. The first term combines the decisions of incumbents and spinout entrants. There are in total $m(s,j)\,(1+\gamma_j)$ such firms making entry decisions, $m(s,j)$ incumbents and $m(s,j)\,\gamma_j$ potential spinout entrants. Since their continuation value is the same, their continuation/entry probability, $\chi_{s,j}$, is also the same. The second item on the right-hand side is the inflow of de novo entrants. Note that the number of de novo potential entrants, n_j, is location specific, and that moreover de novo entrants at location j are ex ante identical in terms of their expected capability.

Given each firm's output level, $q(s; j, p)$, and given the mass of each firm's type, $m(s, j)$, we determine total supply in location j. Aggregating over locations, we get total supply. We assume industry demand is given by the inverse demand function $p = D^{-1}(Q)$. Industry price then clears the market in each period so that total supply equal total demand:

$$p = D^{-1}\left(\sum_{s,j} q(s; j, p)\, m(s, j)\right) \tag{4}$$

■ **Industry equilibrium.** An industry equilibrium is defined by a sequence of prices $\bar{p}^*$, a mass of entrants n^*_{jt}, a measure of incumbent firms $m^*(s, j, t)$, and a policy function $\chi^*(s, j, t)$ such that

- n^*_{jt} satisfies the entry condition for de novo entrants each period, that is, n^*_{jt} satisfies (1);
- $m^*(s, j, t+1)$ is defined recursively given $m^*(s, j, t)$, n^*_{jt}, and $\chi^*(s, j, t)$, according to (2).
- $\chi^*(s, j, t)$ solves incumbent firms and potential spinouts' dynamic optimization problem each period, given their belief of $\bar{p}^*$, that is, $\chi^*(s, j, t)$ satisfies (3);
- p^*_t clears product market each period, that is, p^*_t satisfies (4);

Although our model introduces some specific features — namely the distinction between de novo and spinout entrants — its basic features are similar (and simpler) than the general framework presented in Hopenhayn (1992). With small changes, the equilibrium existence and uniqueness results in Hopenhayn (1992) can therefore be applied in the present context.

■ **Equilibrium properties.** The theoretical model presented above implies a series of equilibrium properties which we now develop formally. (All proofs may be found in the Appendix.)

Proposition 1. *An incumbent (a potential spinout) is more likely to survive (enter) if it belongs to a higher capability family, given the same location and time.*

Proposition 2. *A high-capability family on average has a bigger family size, given the same location and time.*

Proposition 3. *Given positive entry and exit in the stationary equilibrium, spinout firms have lower probability of exit than de novo firms, given the same location j.*

All of these results are consistent with the empirical evidence presented in Section 3. For example, Table 5 shows that family top has a positive effect on spinout rates, while Table 6 shows that family top has a negative impact on exit rates (cf Proposition 1). Propositions 1 and 2 together imply that family size is positively correlated with the spinout rate but negatively correlated with the exit rate, which are consistent with our findings in Tables 5 and 6. Moreover, Tables 6 and 7 show that spinouts have a lower exit rate than de novo firms (cf Proposition 3).

This correspondence between theory and empirical observation gives us confidence in the model as a good description of the evolution of the automobile industry. We next attempt to calibrate the model's parameters with a view at going beyond qualitative description. Specifically, our goal is to use the calibrated model to estimate the relative contribution of each of the model's features, an exercise we refer to as "agglomeration accounting."

■ **Model calibration.** We now calibrate the model so as to quantify the relative importance of each of the agglomeration factors. The reduced-form regressions in the previous section establish three important facts: First, the effect of Marshallian economies is very small (or maybe even negative). Second, location specific effects seem to play an important role, especially regarding Detroit. Third, spinouts play an important role, though this role varies across locations (Detroit being quite idiosyncratic) and across different stages of industry evolution (cf Figure 2).

Consistent with these stylized facts, we developed a model with no intra-industry effects. We are thus left with three possible sources of agglomeration economies: inter-industry effects, spinouts, and location specific effects.

■ **Functional forms.** In the model calibration, we assume six production locations ($j = 1, ..., 6$), corresponding to St. Louis, Chicago, New York, Indianapolis, Detroit, and others;[23] and two levels of firm capability ($s = 1, 2$), corresponding to low and high. We exclude any location fixed effects in production functions so as to limit the number of free parameters, and also to focus on more interesting (and more meaningful) explanatory factors. However, we allow for location-specific effects through differences in entry and spinout rates.

We specify the profit function $\pi(s; p)$ by assuming a decreasing returns production function

$$q(s) = \exp(c_1\, s)\, l^\alpha$$

where c_1 captures the relative advantage of firms with a higher capability, and l is the quantity of input. This implies that a firm, taking industry price p as given, has profit and output given by

$$\pi(s; p) = \left(\frac{1-\alpha}{\alpha}\right) (\alpha\, p)^{\frac{1}{1-\alpha}} \left(\exp(c_1\, s)\right)^{\frac{1}{1-\alpha}}$$

$$q^*(s; p) = (\alpha\, p)^{\frac{\alpha}{1-\alpha}} \left(\exp(c_1\, s)\right)^{\frac{1}{1-\alpha}}$$

Furthermore, we assume that the outside option follows an i.i.d. exponential distribution with parameter σ.

■ **Demand curve.** To calibrate our model, we first estimate an industry demand function using historical annual data of auto prices and output from Thomas (1977). The data range is 1900–1929, and we assume a simple log-log per capita demand function:

$$\log\left(\frac{Q_t}{pop_t}\right) = a_t - b\, \log(p_t)$$

23. The category "others" includes 30 other states. Due to data limitation, we combine New York City and Rochester into one center (New York).

In the regression, we control for log U.S. GDP per capita (as a proxy for income) in the demand intercept a_t. Both auto price and GDP per capita are in real terms.

To address the issue of potential endogeneity of the price variable, we exploit the model structure. In our theoretical model, industry long-run capability distribution is correlated with price, yet uncorrelated with any transitory demand shock. One proxy for long-run industry capability distribution is the share of spinout firms, which we used as an instrumental variable to estimate the demand slope parameter b. Our IV estimation gives $b = 3.39$ (0.39), with standard error in parentheses. The demand shifter is given by $a_t = 0.04 \times \log(\text{GDP per capita})_t + 17.40$.

The first-stage regression results (adj. $R^2 = 0.86$) are given by:

$$\log(p_t) = \underset{(1.28)}{2.86} + \underset{(0.76)}{1.72} \times \log\left(\frac{GDP_t}{pop_t}\right) - \underset{(0.66)}{5.89} \times (\text{Spinout Share})_t$$

The second-stage regression results (adj. $R^2 = 0.83$) are in turn given by:

$$\log\left(\frac{Q_t}{pop_t}\right) = \underset{(5.55)}{17.40} + \underset{(2.17)}{0.04} \times \log\left(\frac{GDP_t}{pop_t}\right) - \underset{(0.39)}{3.39} \times \log(p_t)$$

where standard errors are reported in parentheses.

■ **Other model primitives.** We observe positive de novo entrants at each location during our sample period. Since there is a single automobile market, the data can only be rationalized by a free entry equilibrium with location-specific entry costs. We thus directly feed into the model data of non-spinout entry rates in each location (as explained by our negative binomial regressions reported in Table 4). They range from 1.0 for St. Louis to 7.5 for New York. We also use the logit regression reported in Table 8 to calibrate the top entry probability, that is, $\mu(s = 2)$.[24] Finally, we set the discount factor at $\beta = 0.925$.

■ **Calibrated parameters.** There are five key model parameters left to be calibrated: the parameter reflecting cost heterogeneity across types c_1, the degree of returns to scale α, the average value of the outside option σ, the Detroit specific birth probability γ_D, and the industry average birth probability γ. Our calibration strategy considers 1919 as an industry equilibrium.[25] We pick parameter values so as to match:

- the distribution of output across the five production centers and other regions[26]
- the firm exit rate at the five production centers and other regions
- the spinout rates at the five production centers and other regions

24. The dependent variable is entry of a non-spinoff firm which was ever a top seller in the auto industry. The sample range is 1895-1915; it includes 560 non-spinoff entrants in 5 production centers and 30 other states. The results show that the relative size of related industries, measured by local carriage & wagon employment over population ratio, significantly raises the chance of being a top non-spinoff entrant. The firm entry year is shown to have a negative effect, and per capita income and the Detroit dummy are not statistically significant.

25. We choose 1919 because it's one of the census years.

26. We group state-level automobile output by the five production centers and 30 other states. Data sources: *Census of the U.S. Manufactures,* 1919.

Table 8
Logit models of top non-spinout entry.
Dependent variable: being a top non-spinout entrant

	Spec 1	Spec 2	Spec 3
Entry year	-0.279*** (0.060)	-0.281*** (0.063)	-0.286*** (0.064)
Per capita income 1900	-0.010** (0.005)	0.013 (0.012)	0.022 (0.016)
C&W employment / population ratio 1904		1.142** (0.474)	1.385** (0.640)
Detroit			0.744 (0.546)
Constant	530.291*** (114.088)	525.903*** (120.587)	533.575*** (121.230)
Log Likelihood	-92.104	-86.946	-85.959
N	560	560	560

Notes: Standard errors in parentheses. Star levels: 10, 5 and 1%.

Based on this exercise, we set c_1=0.33, α=0.9, σ=0.86, γ_D=0.08, and γ=0.02. The calibrated value of σ, the average value of the outside option, captures the industry demand and production technology in 1919. In the Appendix, we show that the model is amenable to "rescaling," a fact that allows us this degree of freedom.

■ **Agglomeration accounting.** Table 9 shows the results from our calibration exercise. The second column shows the data, whereas the third column shows the basic model fit. Following that, we have 4 columns which correspond to different counterfactuals. In order to judge the goodness of fit, we present several basic variables of interest related to agglomeration accounting:

- Each region's output share
- The HHI index applied to each region's output share, that is, $HHI = \sum s_i^2$, where s_i is region i's output share
- The prediction mean squared error

Additional variables of interest include

- Exit rates (Detroit and Non-Detroit)
- Spinout rates (Detroit and Non-Detroit)

Finally, in order to account for the effect of each model feature, we run alternative counterfactuals where that feature is shut off. This corresponds to the last 4 columns in Table 9. In Counterfactual 1, we set the variable "Carriage and wagon industry employment" to have the same value in all regions (specifically, the average) when feeding into the model with estimated number and quality of non-spinoff entrants. We thus effectively shut off the type of related industries effects described in Jacobs (1969). In Counterfactual 2, we omit location fixed effects on non-spinout firm entry rates, effectively shutting off the type of location effects described in Ellison and Glaezer (1998). In Counterfactual 3 we force spinout

Table 9
Model fit under basic version and various counterfactuals: 1. Uniform C&W employment levels; 2. No location fixed effects; 3. No spinouts; 4. No Detroit-specific spinouts.

			Counterfactuals			
	Data	Model	1	2	3	4
Output share						
Chicago	0.072	0.035	0.023	0.044	0.062	0.061
Detroit	0.678	0.676	0.179	0.520	0.384	0.385
New York	0.104	0.114	0.072	0.189	0.215	0.214
Indianapolis	0.045	0.083	0.010	0.112	0.167	0.168
St Louis	0.026	0.006	0.011	0.010	0.010	0.009
Others	0.075	0.086	0.704	0.127	0.164	0.163
HHI	0.479	0.479	0.055	0.321	0.226	0.227
Prediction MSE		0.001	0.108	0.007	0.020	0.020
Exit rate						
Non-Detroit	0.150	0.146	0.145	0.144	0.142	0.143
Detroit	0.130	0.135	0.128	0.132	0.135	0.136
Additional moments						
Average center spinout rate	0.030	0.039	0.022	0.032	0.000	0.017
Detroit top spinout rate	0.079	0.073	0.074	0.073	0.000	0.019

rates to be zero in all regions, whereas in Counterfactual 4 we allow for spinout rates but force the Detroit spinout rate to be the same as in other regions, effectively shutting off the type of spinout effects described in Klepper (2007).

Since our goal is to explain agglomeration, it seems natural to use the HHI index of region output shares as an indicator of fit (We will also compare output shares in individual locations). Comparing the data column to the model column, we see that the values of HHI are virtually identical, that is, our model does a very good job at explaining the overall level of industry agglomeration. The low level of the prediction mean squared error reinforces this idea.

Consider now Counterfactual 1. As mentioned earlier, this corresponds to forcing the values of C&W employment to be uniform across regions, thus shutting off the Jacobs (1969) related industry effects. The value of HHI drops from .479 to .055, whereas the prediction MSE increases from .001 to .108. Moreover, the total output share of the five auto production centers falls dramatically, from 91% to 30%. This suggests that the related-industry effect, explained by the variable C&W employment, plays a very important role in explaining the formation of auto production centers.

Omitting region fixed effects, which we do in Counterfactual 2, also implies a lower value of HHI. However, the decrease in HHI is smaller, just as the increase in prediction MSE is also smaller. Note that the total output share of the five production centers falls only a little bit, from 91% to 87%, though the shares of Detroit and New York change quite a bit. Overall, this suggests that region fixed effects account for some agglomeration, but the magnitude is relatively small.

Turning to the effect of spinouts, we first force all spinout rates to be zero. This corresponds to Counterfactual 3, where we see a significant drop in HHI, from .479 to .225. Meanwhile, the prediction MSE increases from .001 to .020. Finally, Counterfactual 4 qualifies the precise channel through which spinouts operate to explain agglomeration. We allow for positive spinout rates but force the rate to be uniform across all regions, whereas in the base case we allow for a different Detroit spinout rate. As can be seen from the last column in Table 9, the drop in HHI is almost identical to that of Counterfactual (3). In other words, the main contribution of spinout rates to explaining industry agglomeration comes from allowing Detroit to have a different spinout rate, an estimate that seems consistent with the work of Klepper (2007).

Before moving on to agglomeration accounting, it is worth mentioning that, in addition to 1919, we also calibrated the model to 1904 data (an earlier census year). One important difference between 1904 and 1919 is that at the early date spinouts were virtually non-existent. In fact, the drop in explained HHI when shutting off spinouts is nearly zero. We conclude that the effect of spinouts varies greatly along the industry life cycle. This is not entirely surprising: by definition, early entrants could not be spinouts, that is, could not possibly have originated from existing industry firms. As a result, the industry development between 1904 and 1919 is similar to comparing Counterfactual (3) or (4) with the base model, which shows that Detroit's output share rises from 38% to 68% as the result of spinouts. To sum up, the agglomeration accounting exercise we present next is based on our 1919 calibration, and for this reason the estimated contribution of spinouts is likely to be over-stated with respect to the overall contribution along the entire industry history.

Based on the results from Table 9, Table 10 presents an estimate of the relative contribution of each of the factors we've considered in the paper. In the second column of Table 10, we note the decrease in HHI that results from omitting a certain feature (as obtained from Table 9). For the first line, intra-industry externalities, the value is obtained by construction, that is, we assume that there are no intra-industry externalities. As mentioned earlier, this follows from the results we obtained in reduced form regressions, where industry variables have very small, or zero, effect on entry and exit rates. For the remaining rows (Location related industries through Spinouts) we compute the change in HHI from Table 9. We exclude Counterfactual 4 from Table 9 because we think that would amount to double counting. However, it should be understood from Table 9 that when we talk about the effect of spinouts we mean primarily the effect of spinouts in Detroit.

There are many ways to compute the relative contribution of each model feature. Table 10 shows three different indicators: the drop in HHI from omitting the particular model feature; the ratio between this change in HHI and the value of HHI in the data; and the ratio between each contribution and the sum total of all contributions. As can be seen from the second indicator, the sum total of all contributions to HHI is about 174%. This suggests that there are important interactions between the various model features, so that a simple additive contribution accounting is subject to possible overestimation of the contribution of

Table 10
Agglomeration accounting: relative contribution of various model components

Factor	Δ HHI	% of HHI	% $\sum \Delta$ HHI
Intra-industry externalities	0.000	00.0	00.0
Location related industries	0.424	88.5	50.8
Location fixed effects	0.158	33.0	18.5
Spinouts	0.253	52.8	30.3
Total	0.835	174.3	100.0

a given feature.[27]

Finally, the last column of Table 10 shows that, in the horse race between the various explanations for industry agglomeration, the development of related industries in each region seems to come out ahead, with a relative contribution of about one-half. Spinouts contribute a little less than one-third, whereas location fixed effects account for a little less than one-fifth. These results are largely consistent with the reduced-form regressions presented in Section 3.

5. Conclusion

Taking the early U.S. automobile industry as an example, we evaluate four competing hypotheses on regional industry agglomeration: intra-industry local externalities, inter-industry externalities, employee spinouts, and location fixed effects. Our findings suggest that inter-industry spillovers, in particular the development of the carriage and wagon industry, play an important role by fostering non-spinout entrants. Spinouts play a secondary role and only contribute to agglomeration at later stages of the industry life cycle. Location fixed effects (beyond the effects through inter-industry spillovers) also explain differences in regional agglomeration, though to a lesser degree. Finally, the presence of other firms in the same industry has a negligible (or maybe even negative) effect on agglomeration.

There are some avenues for further research. First, our paper focuses on the evolution of an individual industry. It would be useful to extend the analysis to comparative studies on multiple industries. Second, due to data limitation, we use entry and exit as proxy measures of firm performance. To the extent that data is available, future studies could use

27. An alternative possibility for agglomeration accounting — in fact, for accounting the relative contribution of any feature to any model — is to adapt the notion of Shapley value: where n was the number of players, think now number of model features; where $v(S)$ was the value of a coalition, think now the model fit when a set S of features is turned on. Then the contribution of feature i to the model's overall fit is given by

$$x_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|\,!\,(n-|S|-1)\,!}{n\,!} \Big(v(S \cup \{i\}) - v(S) \Big)$$

Recall that this is basically the value added by feature i averaged over all possible sets of features S to which feature i is added.

The Shapley value has various desirable features, including that $\sum x_i = v(N) - v(\emptyset)$. This avoids the problem we face in Table 10, whereby the sum of contributions adds to more than 100%.

more direct measures of firm performance, such as output, profit, employment, or product variety. Third, our paper points to the channels through which local externalities and spinouts contribute to the industry agglomeration. It would be interesting to study more precisely the nature and size of local spillovers through those channels. Finally, it would be useful to conduct cross-country comparisons of industry agglomeration in both advanced and developing economies. This will help us better understand the nature of increasing-return technologies and regional spillovers, which are important driving forces for economic growth, development, and international trade.

Appendix: proofs

Proof of Proposition 1: Note that an incumbent's continuing decision is equivalent to a spinout's entry decision. Given that $\pi(s;j,p)$ is strictly increasing in s, continuous, and bounded, standard dynamic programming argument shows that $VC(s;j,\bar{p})$ is continuous in s and strictly increasing in s for $\bar{p} > 0$. Thus we know that for each period, $F(VC(s;j,\bar{p}))$ is strictly increasing in s, given the same location j. ■

Proof of Proposition 2: Note that for each period, all incumbents at location j have the same probability γ_j of having a potential spinout. Also, we show above that an incumbent (a potential spinout) is more likely to survive (enter) if it belongs to a higher capability family. Thus, a higher-capability family has a bigger family size on average, given that $(1+\gamma_j)\chi_{s,j}$ increases in s. ■

Proof of Proposition 3: The stationary distribution is defined by $m_j^* = m_j^*(1+\gamma_j)\chi_j^* + n_j\mu_j$, so $m_j^* = \frac{n_j}{1-(1+\gamma_j)\chi_j^*}\mu_j$. The distribution of spinout firms is $m_j^*\chi_j^* = \frac{n_j\chi_j^*}{1-(1+\gamma)\chi_j^*}\mu_j$. Since χ_j^* is strictly increasing in s, the capability distribution of spinout firms strictly dominates that of *de novo* firms, which is μ_j. ■

Appendix: a note on rescaling

This note proves that the industry equilibrium derived from our model is invariant with respect to a rescaling of the model parameters, such as market size a and production technology $q(s)$, as long as firms' average value of outside option σ is rescaled appropriately.

Given the demand function estimated in Section 4,

$$\log\left(\frac{Q}{pop}\right) = a - b\,\log(p)$$

letting $pop \times \exp(a) = Z$, we can rewrite the market equilibrium condition in steady state as

$$\log\left(\sum_s q^*(s,p)\,m(s)\right) = \log(Z) - b\,\log(p)$$

where $q^*(s,p)$ is the firm's profit-maximizing output given the price p. Profit maximization implies

$$q^*(s,p) = (\alpha\,p)^{\frac{\alpha}{1-\alpha}}\Big(\exp(c_1\,s)\Big)^{\frac{1}{1-\alpha}}$$

Thus the market equilibrium price $p^*(m,Z)$ is determined by

$$\log(p^*) = \frac{\log(Z)}{\frac{\alpha}{1-\alpha}+b} - \frac{1}{\frac{\alpha}{1-\alpha}+b}\;\log\left(\sum_s(\alpha)^{\frac{\alpha}{1-\alpha}}\big(\exp(c_1\,s)\big)^{\frac{1}{1-\alpha}}\,m(s)\right)$$

We can then always compute an industry equilibrium which is invariant to the scale of Z as long as the firms' outside options are subject to rescaling.[28] Let the rescaled price

$$\log(\tilde{p}) = \log(p^*) - \frac{\log(Z)}{\frac{\alpha}{1-\alpha}+b}$$

We can similarly rescale profit $\pi(s,p)$, which is again Z dependent:

$$\pi(s,p^*) = \left(\frac{1}{\alpha}-1\right)(\alpha p^*)^{\frac{1}{1-\alpha}}\Big(\exp(c_1\,s)\Big)^{\frac{1}{1-\alpha}}$$

The profit function is given by $\pi(s,p^*) = Z^{\frac{1}{\alpha+b(1-\alpha)}}\pi(s,\tilde{p})$.

Finally, recall the firm value in steady state is given by

$$VC(s) = \beta\bigg(\pi(s) + F\big(VC(s)\big)\;VC(s) + \Big(1 - F\big(VC(s)\big)\Big)\;E\Big(\phi'|\phi' \geq VC(s)\Big)\bigg)$$

Assuming ϕ is exponentially distributed (that is, F is exponential), we have

$$E\Big(\phi'|\phi' \geq VC(s)\Big) = VC(s) + \sigma$$

28. We can also allow production technology $q(s)$ to be rescaled, e.g. changing $q^*(s,p)$ to $\lambda q^*(s,p)$ in the above equation. We can then define a new scaling parameter $\hat{Z} = Z/\lambda$ and all the following proof goes through.

where σ is the mean of the firm's outside option. Then we have

$$\begin{aligned} VC(s) &= \beta\Big(\pi(s) + \Big(1 - F\big(VC(s)\big)\Big)\,\sigma + VC(s)\Big) \\ &= \beta\Big(\pi(s) + \exp\big(-VC(s)/\sigma\big)\,\sigma + VC(s)\Big) \end{aligned}$$

This implies that if the mean outside option also has a scale factor Z such that $\sigma = Z^{\frac{1}{\alpha+b(1-\alpha)}}\,\tilde{\sigma}$, where $\tilde{\sigma}$ is a constant, then we can go from

$$\widetilde{VC} = \Big(Z^{\frac{1}{\alpha+b(1-\alpha)}}\,\tilde{\sigma}\Big)^{-1} VC$$

to

$$\widetilde{VC}(s) = \beta\Big(\tilde{\pi}(s)/\tilde{\sigma} + \exp\Big(-\widetilde{VC}(s)\Big) + \widetilde{VC}(s)\Big)$$

Hence, the firm's rescaled value function, $\widetilde{VC}(s)$, is invariant with respect to Z. Because firm exit rate is simply $\exp(-\widetilde{VC}(s))$, firm entry and exit rates in the steady state are invariant with respect to Z.

We have thus proved that industry equilibrium is invariant with respect to a rescaling of the model parameters governing market size and production technology, as long as firms' average value of outside option is rescaled appropriately.

References

Arrow KJ (1962), "The Economic Implications of Learning by Doing," *Review of Economic Studies* **29**, 155-172.

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

Buenstorf, Guido, and Steven Klepper (2009), "Heritage and Agglomeration: The Akron Tyre Cluster Revisited," *The Economic Journal* **119**, 705–733.

Cabral, Luís, and Zhu Wang (2009), "Spinouts: Theory and Evidence from Early U.S. Automobile Industry," Federal Reserve Bank of Kansas City Working Paper.

Chatterjee, Satyajit, and Esteban Rossi-Hansberg (2012), "Spin-Offs and the Market for Ideas," *International Economic Review* **53** (1), 53-93.

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, Glenn, and Edward L. Glaeser (1997), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy* **105** (5), 889–927.

Ellison, Glenn, and Edward L. Glaeser (1999), "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration," *American Economic Review* **89** (2), 311–316.

Ellison, G. and E.L. Glaeser and W. Kerr (2010), "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review* **100** (3), 1195-1213.

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

Glaeser, Edward L., Hedi D. Kallal, José A. Sheinkman, and Andrei Schleifer (1992), "Growth in Cities," *Journal of Political Economy* **100**, 1126–1152.

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

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

Henderson, Vernon, Ari Kuncoro, and Matt Turner (1995), "Industrial Development in Cities," *Journal of Political Economy* **103**, 1067–1090.

Hopenhayn, Hugo (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica* **60** (5), 1127–1150.

Jackson, John (1988), "Michigan," in R. Scott Fosler (Ed), *The New Role of American States*, New York: Oxford University Press.

Jacobs, Jane (1969), *The Economy of Cities*, New York: Vintage Books.

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

Klepper, Steven (2007), "Disagreements, Spinouts, and the Evolution of Detroit as the Capital of the U.S. Automobile Industry," *Management Science* **53** (4), 616–631.

Klepper, Steven (2010), "The Origin and Growth of Industry Clusters: The Making of Silicon Valley and Detroit," *Journal of Urban Economics* **67**, 15–32.

Krugman, Paul (1991), *Geography and Trade*, Cambridge: MIT Press.

Marshall, Alfred (1890), *Principles of Economics*, Macmillan, London.

Romer, Paul M. (1986), "Increasing Returns and Long-run Growth," *Journal of Political Economy* **94**, 1002-1037.

Rosental, Stuart and William Strange (2004), "Evidence on the Nature and Sources of Agglomeration Economies," *Handbook of Urban and Regional Economics* **4**, 2119–2171.

Rubenstein, James M. (1992), *The Changing US Auto Industry: A Geographical Analysis*, New York: Routledge.

Smith, Philip H. (1970), *Wheels within Wheels: A Short History of American Motor Car Manufacturing*, New York: Funk and Wagnalls.

Thomas, Robert Paul (1977), *An Analysis of the Pattern of Growth of the Automobile Industry 1895-1929*, New York: Arno Press.