THE UNIVERSITY OF WARWICK

June 2015 No.235

# Agglomeration Economies and Productivity Growth : U.S. Cities, 1880-1930

**Alexander Klein & Nicholas Crafts**

## WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

# AGGLOMERATION ECONOMIES AND PRODUCTIVITY GROWTH: U.S. CITIES, 1880-1930


**Alexander Klein**

University of Kent

and

**Nicholas Crafts**

University of Warwick


June 2015


**Abstract**

We investigate the role of industrial structure in labor productivity growth in U.S. cities between 1880 and 1930 using a new dataset constructed from the Census of Manufactures. We find that increases in specialization were associated with faster productivity growth but that diversity only had positive effects on productivity performance in large cities. We interpret our results as providing strong support for the importance of Marshallian externalities. Industrial specialization increased considerably in U.S. cities in the early 20th century, probably as a result of improved transportation, and we estimate that this resulted in significant gains in labor productivity.



**Keywords**: agglomeration economies; Jacobian externalities; manufacturing productivity; Marshallian externalities; industrial structure

**JEL Classification**: N91; N92; O7; R32

**Acknowledgements**: we have benefited from comments made by participants in seminars at Stanford, UC Davis, UC Irvine, Brunel University, and Cardiff Business School as well as by Dan Bogart, Chris Meissner, and Gavin Wright. The usual disclaimer applies.

## I. Introduction

In the last 20 years or so there has been a great revival of interest in the nature and extent of agglomeration economies that cities generate. It is widely recognized that making a success of urbanization is fundamental to the achieving economic development (Duranton, 2014). The study of agglomeration externalities has developed so rapidly under the auspices of the new economic geography that major survey articles are frequently written (Rosenthal and Strange, 2004; Combes and Gobillon, 2014). Yet, remarkably little is known about the extent or nature of such externalities in U.S. cities during American industrialization (Kim and Margo, 2004). This is a gap that deserves to be filled especially given that cities were growing rapidly around the time at which the United States overtook the United Kingdom to become the world's leading economy during the so-called 'second industrial revolution'. Between 1880 and 1920, the percentage of urban population rose from 28.2 to 51.2. Cities became larger; in 1880 there were 41 cities with population of 50,000 or more accounting for 14.3 per cent of the population but by 1920 this had risen to 144 cities with 31.0 per cent of the population. This became the era of large industrial cities (Kim, 2000). Cities also became more specialized as is epitomized by the examples of Akron, Ohio and Detroit.

The key questions that we address in this paper based on a detailed analysis of data in the Census of Manufactures are the following

1) Did Marshallian and/or Jacobian agglomeration externalities raise labor productivity growth in American cities in the period between 1880 and 1930?
2) Did productivity increases associated with changes in the pattern of American urbanization make a substantial contribution to American economic growth at the start of the 20$^{th}$ century?

Our empirical strategy is to run regressions that relate labor productivity growth in manufacturing at the city-industry level during periods of between 20 and 50 years between 1880 and 1930 to measures of the degree of industrial specialization and diversity of urban economic activity while controlling for a wide range of variables relating to geography and state-level factor endowments. We have constructed a unique dataset derived from the Census of Manufactures for this purpose.

The main contribution of the paper is to establish the connection between specialization and diversity of industrial structure in early-20$^{th}$ century American cities and labor productivity growth. In particular, we report four main findings. First, we show that greater specialization was correlated with faster labor productivity growth in a city's industrial sectors. This effect is both economically and statistically significant. Second, we find that greater diversity of a city's industrial structure has a non-linear effect on productivity growth. For small cities, the effect is negative but at larger city size (bigger than Cleveland) the effect is positive. Third, taking these first two findings together, we interpret the results as strong evidence of Marshallian (intra-sector) agglomeration externalities but rather weak evidence of the importance of Jacobian (inter-sectoral) externalities. Fourth, we note that specialization in our sample of cities increased on average by about 25 per cent on our measure and that our regressions indicate that this would have had a substantial impact on labor productivity, raising it by about 23 per cent between 1890 and 1920. A comparison with the estimates in Kendrick (1961) indicates that these productivity gains from increased specialization in U.S. cities accounted for over half of all manufacturing productivity growth in the period.

These results imply that the United States did indeed make a success of urbanization during the second industrial revolution. A reasonable conjecture based on the traditional economic geography literature is that the increased specialization of cities from the late 19th century was predicated on improvements to the transport system (Pred, 1966). This would potentially represent an important addition to the estimates of the 'social savings' of railroads made by Fogel (1964).

The paper proceeds as follows. Section II presents a detailed account of the literature on agglomeration externalities as it relates to our research. Section III sets out our approach to estimation and Section IV describes the data. In Section V, we present detailed results including an analysis of the economic significance of the econometric estimates and we provide a review of some wider implications of the results in Section VI. Section VII concludes.

## II. Related Literature

Agglomeration externalities have received a good deal of attention since the advent of new economic geography. Good overviews of the basic ideas and how they can be modeled can be found in Duranton and Puga (2004) (2013). The general idea as expressed in a production-function format is that

$$Y_{it} = G_{it}(.)F_{it}(X_{it})$$

where Y is output, X is a composite input, and G(.) is a shift factor which raises the productivity of this input. Agglomeration is seen as the underpinning for the shift factor

The sources of agglomeration externalities can be categorized as sharing, matching and learning mechanisms. Sharing mechanisms include sharing indivisible facilities, sharing the gains from a wider variety of input suppliers, sharing the gains from a finer division of labor, or sharing risks. Matching mechanisms include better expected quality and/or higher probability of matches especially between employers and workers in a larger labor market. Learning mechanisms relate to the diffusion and accumulation of knowledge including tacit knowledge which is enhanced by the proximity of other producers. These benefits of agglomeration may occur within industries or across sectors. The former can be thought of as localization externalities which accrue from specialization and the latter as urbanization externalities which stem from diversity of production. Conventionally, external economies from specialization are described as Marshallian (Marshall, 1890) while those from diversity are termed Jacobian (Jacobs, 1969).

The general idea of an agglomeration externality which accrues within a sector can be illustrated on the basis of a model with a sharing mechanism which has the properties that an increase in the labor input of a sector is associated with more intermediate producers and that final goods producers become more productive when they have access of a wider range of intermediate goods (Duranton and Puga, 2004). This will result in an equation for aggregate production in a sector as follows:

$$Y = \left[n(x)^{1/(1+\varepsilon)}\right]^{1+\varepsilon} = L^{1+\varepsilon}$$

where n is the number of intermediate producers, x is the aggregate amount of intermediates used, L is total labor supply in the sector, and ε is elasticity of substitution. This exhibits increasing returns to scale at the sector level.

Extending the argument to see how the production externality can become not only an agglomeration force but also an engine of growth when workers who accumulate human capital find it worthwhile to agglomerate in larger cities and the agglomeration economies that ensue amplify the effect of human capital accumulation on growth, consider the following production function for city output (Duranton and Puga, 2013)

$$y = \beta H^{\sigma} h^{\alpha} l^{1-\alpha}$$

where y is output per worker, H = hN, i.e., human capital per worker times the number of workers, $\beta H^{\sigma}$ is a city-level externality based on aggregate human capital, h is human capital per worker and l is labor time per worker. The evolution of output per worker based on human capital accumulation will be

$$(y_t/y_{t-1}) = (h_t/h_{t-1})^{\alpha+\sigma}(N_t/N_{t-1})^{\sigma}$$

so positive agglomeration economies ($\sigma > 0$) raise the growth rate.

An alternative formulation which also allows externalities from human capital formation to drive growth was provided by Lucas (1988) who speculated that knowledge spillovers in the form of Jacobian externalities from the growth of diverse cities might be the engine of growth in the 20th-century United States. The production function might then be written as

$$Y = AK^{\gamma}(hL)^{1-\gamma}h_e^{\varphi}$$

where $h_e^{\varphi}$ is the external effect of human capital which performs a role similar to exogenous technological progress in the traditional neoclassical model.

The production functions in the above discussion are intended simply to illustrate how agglomeration economies might affect output growth rather than as models to be estimated. In fact, the empirical literature on agglomeration economies offers several estimation strategies (Combes and Gobillon, 2014) and, as discussed in the next section, we choose an empirical specification which links labor productivity to measures of Marshallian and Jacobian externalities and which models these externalities as contributing to the Hicks-neutral factor-augmenting technology level.

While cities can generate agglomeration economies at the same time increasing city size tends to raise congestion costs. Thus efficient city size decreases as commuting costs increase. If diversity does not offer agglomeration economies but specialization does, then diversity will undermine the realization of total agglomeration externalities by raising commuting costs for workers. Then, if final goods can be traded costlessly between cities, we would expect to see completely specialized cities whose size would depend on the extent of Marshallian externalities in their industry (Henderson, 1974). More generally, as transport costs decline and more final goods can be imported into a city, additional gains from increased specialization are possible

(Abdel-Rahman, 1996). So, improvements in transportation which reduce the costs of commuting and trading tend to facilitate Marshallian externalities by permitting larger and more specialized cities.

A good deal of empirical evidence on the impact of city size on productivity and on the nature of agglomeration externalities is available for the recent past and this has been the subject of two comprehensive survey articles by Rosenthal and Strange (2004) and by Combes and Gobillon (2014). The general thrust of this literature can be summarized as follows.

There is general agreement that increases in city size generally raise productivity (or wages as a proxy), although recent studies which control better for sorting effects tend to find lower elasticities. Combes and Gobillon (2014) suggest an elasticity of TFP to population density of about 0.035 to 0.040 might be reasonable. Evidence for the late 20th century generally gives quite strong support for the existence and importance of Marshallian externalities. For example, this result is obtained by Cingano and Schivardi (2004) for Italy, Harris and Moffat (2012) for the United Kingdom, and Henderson (2003) and Henderson et al. (1995) for the United States. Evidence for Jacobian externalities is more mixed and the strongest results are that positive effects of diversity are found for young hi-tech industries as in Harris and Moffat (2012) and Henderson et al. (1995). However, Glaeser et al. (1992) found that dynamic Jacobian externalities were important in the United States in that industrial employment growth was encouraged by diversity in American cities in the period 1956 through 1987.

When we turn to the era of the late 19th and the early 20th century, evidence about agglomeration economies is quite thin on the ground, as is evident from the survey article by Kim and Margo (2004). Kim (2006) showed that urban manufacturing firms had substantially higher wages, labor productivity and TFP than rural firms and that the wage premium increased considerably between 1850 and 1880. On the other hand, Bostic et al. (1997) found no evidence that variables designed to capture ‘urbanization’, ‘specialization’ or ‘localization’ had a positive effect on labor productivity growth in American cities in the 1880s. Pred (1966) provided an interpretation of urban growth in the period as strongly influenced by declining transport costs which facilitated increased city specialization and likewise Chandler (1977) memorably saw the late 19th century as the point at which integration of the domestic market underpinned the beginnings of mass production together with mass distribution. No-one, however, has provided a serious quantification of the extent or nature of agglomeration economies in the early 20th century.

## III. Estimation Framework

We use a simple model of city growth which incorporates local spillovers within and across industries to motivate the empirical framework discussed toward the end of this section. Specifically, we put the framework of dynamic agglomeration spillovers introduced by Glaeser et al (1992) into the simple model of Combes et al (2008) which we have chosen because it matches our data best as it operates with value-added per worker.[1] We consider an industry of sector *s*, in a city *c*, at time *t* with a Cobb-Douglas production function

[1] We do not use total factor productivity as our dependent variable because (i) US Census of Manufactures in 1930 did not report estimates of city-industry capital stock, and (ii) city-industry capital stock figures in 1880-1920 are of very low quality.

$$y_{cst} = A_{cst} l_{cst}^{\mu} k_{cst}^{1-\mu} \quad (1)$$

where $A_{cst}$ is the Hicks-neutral factor-augmenting technology level, $l_{cst}$ is labor and $k_{cst}$ is the quantity of other inputs. The profit is given by

$$\Pi_{cst} = \sum_{c} p_{cst} y_{cst} - w_{st} l_{cst} - r_{st} k_{cst} \ (2)$$

where $y_{cst}$ is the quantity exported to city c, $p_{cst}$ is the mill price in city *c* net of marginal costs of the intermediate inputs, $w_{st}$ is the wage rate and $r_{st}$ is the price of intermediate inputs other than labor. The profit function can be rewritten as

$$\Pi_{cst} = p_{st} y_{st} - w_{st} l_{cst} - r_{st} k_{cst} \ (3)$$

where

$$p_{st} = \sum_{c} p_{cst} \frac{y_{cst}}{y_{st}} \qquad (4)$$

is the average unit value of the produced goods net of the cost of intermediate inputs. As a result, $p_{st}$ $y_{st}$ is sector's *s* value added. The standard first order conditions plugged into the equation 1 yield the average labor productivity defined as value-added per worker

$$\frac{p_{st} y_{st}}{l_{cst}} = (1-\mu)^{(1-\mu/\mu)} \left( \frac{p_{st}^{1-\mu} A_{cst}}{r_{st}^{1-\mu}} \right)^{1/\mu} \quad (5).$$

Following the dynamic externalities framework of Glaeser et al (1992), the growth rate of technology is given by

$$ln \left( \frac{A_{cst+m}}{A_{cst}} \right) \equiv E_{cst} \qquad (6)$$

where $E_{cst}$ is the amount of localized spillovers in industry *s* in a city *c*. Theory does not offer us any concrete functional form linking the growth rate of technology with agglomeration externalities, hence, at this stage, we postulate that $E_{cst}$ is a linear function of Marshallian (Marshall-Arrow-Romer) externalities $MAR_{cst}$ and Jacobian externalities $JACOBS_{cst}$. We can use equation 6 to derive the growth rate of city-industry value-added per worker as a function of the growth of technology. Specifically, we transform equation 5 in logs, express it for period *t+m*, plug both log-equations for *t* and *t+m* into equation 6 to obtain

$$ln \frac{vapw_{cst+m}}{vapw_{cst}} = \frac{1}{\mu} E_{cst} + \left( \frac{1-\mu}{\mu} \right) \left[ \left( ln \frac{p_{st+m}}{p_{st}} \right) - \left( ln \frac{r_{st+m}}{r_{st}} \right) \right] \qquad (7)$$

where, for the sake of clarity in equation (7), we abbreviate value-added per worker as *vapw*. This equation expresses the growth rate of value-added per worker in terms of (i) localized externalities which depends of MAR and Jacobian externalities, and (ii) the difference between the output prices and the prices of inputs. It motivates the regression equation.[2]

Before specifying the regression equation, we define Marshallian and Jacobian externalities. We closely follow Combes (2000) and Cingano and Schivardi (2004) who carefully discuss how to measure these externalities. Marshallian externalities are captured by the degree of specialization

$$MAR_{cs} \equiv (L_{cs}/L_c) \qquad (8)$$

where $L_{c,s}$ is employment in sector $s$ and city $c$, and $L_c$ is the total employment in city $c$. Jacobian externalities are captured by the degree of industrial sector variety outside sector $s$ in the city which is expressed using a Hirschman-Herfindahl index and defined as:

$$JACOBS_{cs} \equiv \sum_{j \neq s} \left(\frac{L_{cj}}{L_c - L_{cs}}\right)^2 \quad (9).$$

The value of the index is 1 if employment outside industry $s$ is concentrated in a single sector, and $1/(Jacobs_c)$ if employment outside industry $s$ is uniformly distributed across other sectors.[3]

We specify the regression equation as follows:

$$ln\frac{vapw_{cst+m}}{vapw_{cst}} = \alpha + \beta_1 MAR_{cst} + \beta_2 JACOBS_{cst} + \beta_3 X_{cst} + \theta_s + \lambda_c + \epsilon_{cs} \quad (10).$$

where $X_{cst}$ is a vector of city-industry controls, $\lambda_c$ is a vector of industry dummies to control for industry-level input and output prices, and $\theta_s$ is a vector of city dummies to control for any unobserved city-effect to avoid omitted variable bias. The error term is expressed as $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$. We estimate regression equation (10) with OLS but there are a couple of econometrics issues to be addressed. First, the errors, $\epsilon_{cs,}$ could be correlated within cities as well as industries. Therefore, we use the two-way clustering method of Cameron et al. (2011). Second, unobserved factors, $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$, may bias the estimates of Marshallian and Jacobian externalities.

Two sorts of biases can be imagined. First, firms might predict $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$ and relocate in anticipation. For example, $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$ may include city-industry specific long-run trends such as changes in technology. If a particular industrial structure is more conducive to technological change, firms might relocate because they predict technological change which would make the industrial composition of cities potentially endogenous. In our case, if, for example, a firm in 1880 can predict technological change between 1880 and 1920, and can predict that this technological change is more likely to happen in highly diversified cities, then it

[2] We would like to stress that we perform not a structural but a reduced-form regression analysis.
[3] In the regression analysis, we use a logarithmic transformation of the diversity index which allows us a very useful interpretation of the change of diversity index as percentage change. The literature uses sometimes the log of diversity index (Combes, 2000) and sometimes an index without logarithmic transformation (Henderson et al., 1995; Cingano and Schivardi, 2004). As a robustness check, we performed the regression analysis without the logarithmic transformation of the diversity index and the results (available from the authors upon request) were qualitatively similar.

will relocate to those cities, making the beginning-of-period industrial structure endogenous, hence biasing our estimates of externalities. Second, there might be other factors which influence Marshallian and Jacobian externalities, and are correlated with the value-added growth rates but which are not controlled for by either $X_{cst}$, industry dummies, or city dummies. We believe that we control adequately for unobserved effects and that the industrial structure of cities at time $t$ is exogenous to the subsequent growth of value-added per worker over the period from $t$ to $t+m$. Nevertheless, as a robustness check, we also estimate equation (10) using instrumental-variables estimation. A thorough discussion of our argument for exogeneity of industrial structure as well as our instruments is provided in the section on instrumental-variables estimation. Equation 10 will be estimated for the periods 1880-1920, 1880-1930, 1890-1920, 1890-1930, and 1900-1920, 1900-1930.[4]

## IV. Data

We created a unique data set of SIC 3-digit city-industry real value added per worker, sectoral specialization, industrial variety, and U.S. state-level employment at SIC 2-digit level in the period 1880-1930. The data come from several sources. City-industry nominal value-added, number of workers, sectoral specialization, and industrial variety are drawn from the U.S. Census of Manufactures and the aggregation of individual industries at the three-digit level follows the standard industrial classification.[5] U.S. state-level employment at SIC 2-digit level in 1880-1920 comes from Klein and Crafts (2012), supplemented by data from the 1930 U.S. Census of Manufactures. The aggregation of individual industries at two-digit level again follows the standard industrial classification. We use the wholesale price index offered by the U.S. Historical Statistics Millennial Edition, series Cc-126 to calculate city-industry real value-added.[6]

The U.S. Census of Manufactures, while an excellent and unique source of city-industry level data, is not without its challenges. First, the 1880 census reports the number of workers and not the number of employees which begins only in 1890. Therefore, we use the number of workers to be consistent over the entire period 1880-1930. Second, value added, as reported by the census in 1900, 1920, and 1930, does not exclude some of the costs.[7] Therefore, we calculated value added by subtracting all the costs from the reported gross value of output. Third, the number of cities included in the censuses changes over time; we deal with this issue in the next section.

Before we present the regression analysis, we provide summary statistics of our data set and various descriptive statistics which offer insights into a relationship between labor productivity and agglomeration economies. Table 1 shows the distribution of the cities in our sample across time and U.S. regions. We see that the 1900 Census of Manufactures contains the most comprehensive information on the U.S. cities. Most of the cities come from the New England,

[4] We do not use the data from the Census of Manufactures in 1910 because its quality is inadequate for the reconstruction of reliable SIC 3-digit level industrial groups.

[5] The assignment of individual industries into each SIC 3-digit level category is available from the authors upon request.

[6] We have also used the wholesale price indices at SIC 2-digit level offered by Cain and Paterson (1981). Their indices, however, run only up to 1919, and so we use them as a robustness check for the regressions of 1880-1920, 1890-1920, and 1900-1920.

[7] Rents for offices are excluded in 1900 and 1920, value of fuel in 1930.

Middle Atlantic and East North Central regions, reflecting the dominant position of those regions in American economic development at the turn of the twentieth century. Table 2 presents the number of matched cities across the periods of our interest. Two points are worth mentioning. First, we see that the number of cities which can be consistently followed over time increased between 1880 and 1900. Second, there is a drop in the number of cities reported in the 1920 and 1930 Census of Manufactures resulting in fewer cities which can be followed up to 1920 and 1930 respectively.

There are two ways to deal with the changing number of cities. We can create a sample of only those cities for which the census consistently provides data in every period we analyze. This would give us a dataset with the same cities throughout. Alternatively, we can use all the cities provided by the census and have a different number of cities in each period we analyze. We do both. Analyzing the latter data set, however, brings up the issue of how to compare the results across different periods. For example, the results for the period 1890-1920 can be different from 1880-1920 because, in addition to the cities which were in the 1880-1920 sample, we have 'new' cities entering the sample of 1890-1920 as well as cities exiting from the 1880-1920 sample.

We do not think that the changing number of cities across periods is the result of sample selection. It is highly unlikely that the changing number of cities had anything to do with unobserved factors which might be correlated with labor productivity growth since decisions to include new cities in the census were driven by the capabilities and financial resources of the Census Bureau. However, it will be useful to see what kind of cities were entering and exiting the sample in the period of 1880-1900 and stayed there until 1920 and 1930, respectively. There was only one town which was present in 1880 but not after, namely, Saint Louise, Maine and two towns were present in 1880 but not in 1900, namely, Saint Louise, Maine, and Washington DC. As for 1890, there are only two towns present in 1890 but not in the 1900 census: Lincoln, Long Island, and Long Island City. The cities which entered the sample in 1890 and 1900, and stayed there until 1920 and/or 1930 were mostly small and medium size towns except for St. Louis, Akron (Ohio), Portland (Oregon), Seattle, and Los Angeles which entered the sample in 1890. As for the degree of industrial specialization and industrial diversity, cities entering the sample in 1890 are not, on average, different from the cities already in the sample. Cities entering the sample in 1900 are slightly more specialized and slightly less diversified. We believe, however, that this will not bias our results since there are only 9 new cities entering the sample for 1900-1920. Nevertheless, we conduct a series of robustness checks in the next section.

As we have seen in Table 2, it is a feature of the Census of Manufactures in 1930 that fewer cities were included in the census relative to the earlier censuses. As a result, our sample of cities in the period ending in 1930 drops relative to the one ending in 1920. Again, this was due to the resources of the US Census Bureau so that it is very unlikely that there would be any kind of systematic relationship between the change in the sample size and the growth of labor productivity of cities and we do not need to worry about sample-selection bias. The towns which dropped out of the sample (but were present until 1920) were mostly smaller towns with an average population of about 40,000, together with two large cities: Omaha and Milwaukee. As for the degree of industrial specialization and diversity, these towns are, on average, more specialized and less diversified than those which remained in the sample. Again, we believe that this will not materially affect the overall results; nevertheless, we conduct a series of robustness checks in the next section.

The period 1880 to 1930 witnessed the 'second industrial revolution' and our data set reflects that. Table 3 presents the average share of employment in selected SIC 3-digit level traditional and modern industries in 1880 and 1930, respectively. We see a clear trend: traditional industries such as textiles (SIC 223 and 225) or leather (SIC 313 and 319) are declining while modern industries such as chemicals (SIC 291 and 287), electrical industries (SIC 362 and 364) and car manufacturers (SIC 371) are rising. Table 4, which shows the fastest- and the slowest-growing industries in U.S. cities, corroborates those results. We see that chemicals, steel, and industries producing instruments and devices were among the fastest growing while food processing, leather, or printing were among the slowest. Furthermore, Table 5 reports the average annual growth rate of value-added per worker aggregated at the SIC 2-digit level and also shows that modern industries are among the fastest growing. Of course, Tables 3-5 present unconditional averages, and the year 1930 reflects the beginnings of the Great Depression which might downward bias the productivity growth figures. Nevertheless, they provide a clear picture of the increasing dominance of the modern industries.

Since this paper focuses on Marshallian and Jacobian agglomeration economies, it is interesting to see which industries were found in the most specialized and the most diversified industrial environments. We take the index of specialization for each city to be the average *MAR* score for its industries and likewise the index of variety to be the average *JACOBS* score for its industries. Table 6 presents the top five and the bottom five industries according to the average specialization and variety indices that they faced across all cities in 1880, 1900, and 1930.

We see in Panel A that traditional industries, such as textiles, were in cities of high specialization, though by 1930, this was also true of modern industries such as petroleum refining and motor vehicles, reflecting increased concentration of those industries in cities (e.g. motor vehicles in Detroit). On the other hand, modern industries such as chemicals and household appliances were located in cities with a diversified industrial structure. This is confirmed when looking at Panel B which shows that modern industries such as plastic materials or household appliances were in cities with the highest variety index while traditional industries such as textiles, food, or the leather industry were in locations with the lowest variety index. It is interesting to see, however, that by 1930, some of the modern industries such as aircraft were located in the cities with a low variety index, indicating that they had become highly specialized.

Figure 1 graphs city populations against their specialization index and to a first approximation this shows that the smallest cities were the most specialized. On a closer look, the relationship, however, is a bit more complex. Indeed, at the lower end of the spectrum, with population up to about 50,000, there were both quite specialized as well as quite diversified towns. For example, in Hamilton, Ohio, population 23,914 in 1900, almost 50 percent of all workers worked in iron and steel foundries; or in South Omaha, Nebraska, population 26,001 in 1900, 89 percent of all workers worked in slaughtering and meat packing industry. On the other hand, there are towns such as Racine, Wisconsin, population 29,102 in 1900, in which no industry was dominant (21 percent in agricultural implements, 19 percent in the production of carriages, wagons, and bicycles, 12 percent in leather, 8 in metal products, and 8 percent in primary metal production). Then, as city population increases, the degree of industrial specialization declines rather rapidly in a curvilinear fashion, until specialization does not vary with the size of the cities.

Figure 2 shows that there seems to be a positive relationship between the city size and industrial diversity. However, that relationship is, again, not linear, but this time seems to be concave. As with Figure 1, up to a population of about 50,000 there were both highly diversified towns as well as more specialized towns. Then, as population increases, cities become more and more diversified.[8] Even so, a clear outlier is Pittsburgh, which, even though it was one of the largest cities with population of 451,512 in 1900, the dominant industry was iron and steel with about 35% of city's employment. Cleveland can also be considered as an outlier with the iron and steel industry being the largest in the city.

Figure 3 offers a comparative perspective on the relationships between city size and specialization, and between city size and variety over the years 1890 through 1920. In the former case, the fitted curve shifts out appreciably after 1900 so that the degree of specialization for a given population was increasing over time. Not surprisingly, in the latter case, the fitted curve shifts in quite considerably so the degree of diversification for a given population was decreasing over time. Detroit is a well-known example of this process, having been a city of diversified industrial composition before becoming a highly specialized city with a dominant car industry by 1920.

A first look at the relationship between labor productivity of manufacturing industries and the specialization and variety indices is provided in Figures 4 and 5, respectively. These figures show that, while the average labor productivity in a city was positively related to the level of industrial specialization of the city, there seems to have been a convex relationship between the index of variety of manufacturing industries and their productivity. This is indeed an interesting relationship and, to our knowledge, it has not previously been observed in the literature. The convexity in Figure 4 implies that to enjoy Jacobian externalities required location in a large city and that the benefits of agglomeration were lower than their costs in the cities in the middle of the diversity spectrum. We shall see in the next section whether this convex relationship also emerges in a regression framework.

## V. Empirical Results

### Basic results

The results of the estimation of equation (10) are presented in Tables 7, 8, and 9. In addition to the main variables of interest — *MAR* and *JACOBS* which capture Marshallian and Jacobian externalities, respectively — the regression controls for city and industry fixed effects, beginning-of-period value-added per worker, and beginning-of-period share of state-level employment in the U.S. total for the corresponding SIC-2 digit level. This last is used to control for the presence of spatially correlated omitted variables operating at the state-level and also at the higher level of industrial aggregation. For example, there might be state-specific effects influencing city-industry productivity not controlled by city-industry dummies, for example, from state-specific factor endowments. Controlling for industry-specific effects at the higher-level industrial classification is important because there might be spillovers operating within

[8] For presentation purposes, Chicago, Philadelphia, and New York City were excluded from Figures 1-2 because their population is more than three times larger than the population of the fourth largest city – St. Louise. Similarly, the city of Spokane was excluded from Figures 4-5. The qualitative features of these figures, however, are not changed. Figures including all cities are available from the authors upon request.

industry groups. For example, labor productivity of industries producing transportation equipment could be influenced by unobserved industry effects specific to the transportation industry; hence our decision to control for industry employment at SIC 2-digit level. We have also estimated equation (10) with city latitude and longitude, access to waterways, lakes, and oceans, distance to New York City, temperature, share of state-level employment in mining, state-level prices, and share of state-level skilled occupations.[9] The results for the main variables of interest were unchanged, reassuring us that our estimates of Marshallian and Jacobian externalities are not affected by geography and state-level factor endowments.

Table 7 shows that average annual labor productivity growth is positively related to *MAR* and negatively to *JACOBS*. The positive Marshallian relationship is statistically significant in all periods but 1890-1930 while the negative Jacobian relationship is significant in all periods but 1880-1930 and 1890-1930. A test of joint significance of specialization and variety is significant in every period but 1880-1930. As for the remaining regressors, the share of state-level SIC 2-digit level employment is not significant whereas value-added per worker at the beginning of the period is significant and negative, indicating a β-type convergence among SIC 3-digit level industries.

Table 8 adds $JACOBS^2$ to the estimating equation to investigate the possibility of a quadratic relationship between average annual labor productivity and Jacobian externalities. The results show that the relationship is convex and that a positive effect is only found for large cities, as suggested by Figure 5. For example, in the estimation for 1890-1920, the city at the lower turning point is Detroit and those to the right of this include Boston, Chicago, New York City, Philadelphia, and San Francisco while those not big enough to enjoy positive Jacobian externalities include Baltimore, Denver, Minneapolis, and Washington DC.

This relationship is not significant in the periods 1880-1930 and 1890-1930, although all agglomeration externalities are jointly significant (in the period 1880-1930 marginally at 13%). We should be cautious about this insignificance since it may reflect failure of the Census to collect information on small and medium-sized towns so that there is insufficient variation in the data capturing the left part of the convex curve. Further investigation of this relationship using a semi-parametric technique which gives some support to positive Jacobian externalities for industrial variety indices greater than 2, which are generally found in relatively large cities, is reported in an appendix.

As we discussed in the previous section, the number of cities changes across periods under consideration. Even though the cities which enter our sample in 1890 and 1900 are not very different from the cities already in the sample, and the results are robust across all six periods, their inclusion might raise an issue of the comparability of the results between 1880, 1890, and 1900. Therefore, Table 9 presents the estimation results when we keep the number of cities constant over time. The first four columns show the results for the periods starting in 1890 and 1900 and only for the cities of the 1880 Census of Manufactures; the last two columns for the periods starting in 1900 and only for the cities of the 1890 Census of Manufactures. We see that the results are consistent with the results in Table 7 and 8, and the only difference is for the period 1890-1920 in which the quadratic term of Jacobian externalities loses conventional

[9] The results are available from the authors upon request.

statistical significance, although the joint significance of the linear and quadratic terms is highly statistically significant.

Overall, it is clear that there is strong evidence that *MAR* has a positive impact on labor productivity growth in manufacturing in our sample of cities during the early 20$^{th}$ century. This result which indicates that Marshallian agglomeration externalities were obtained is consistent with the findings of the literature on the late 20$^{th}$ century reviewed above. For the 1890-1920 sample, the difference in predicted labor productivity growth from Marshallian externalities between the 25$^{th}$ and the 75$^{th}$ percentile city is 0.09 per cent per year. The evidence on Jacobian externalities is, however, altogether more mixed, as one might also expect. On balance, we believe that there is some support for a positive impact of industrial diversity at high values of *JACOBS* but that this was most likely to be found in large cities whereas in smaller cities it is likely that diversity had a negative impact on productivity growth. For the 1890-1920 sample, the difference in predicted labor productivity growth from Jacobian externalities between the 25$^{th}$ and 75$^{th}$ percentile city is 0.119 per cent per year (from -0.086 to + 0.033).

**Instrumental variable estimation**

As was discussed earlier, there is a possibility that, even after controlling for industry- and city-specific effects as well as state-level employment structure, the beginning-of-period industrial structure of cities is still correlated with unobserved city-industry effects, $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$. For OLS to be consistent there must be no link between factors which affect the initial industrial structure of cities and general improvements that affect the productivity of industries in the decades to come. We believe that this was indeed the case as research on technological progress in the United States in the 'second industrial revolution' shows.

Technological progress accelerated at this time but its progress was quite erratic and the development of new technologies was unpredictable. It took 40 years after the commercial generation of electricity to work out how to transform many American factories through electrification in the 1920s (David, 1991). Automobile registrations rose from 8000 in 1900 to 23 million in 1930 but at the start of the 20$^{th}$ century the future had appeared to be steam-powered cars rather than the internal combustion engine. Estimates of TFP growth in Kendrick (1961) are 2.0 per cent per year in the 1880s, 1.1 per cent in the 1890s, 0.7 per cent in the 1900s, 0.3 per cent per year in the 1910s, and 5.3 per cent in the 1920s. New industries (entertainment, electric utilities, and transport equipment) averaged the highest TFP growth rates over these decades but, at the industry level, rank correlation coefficients between decadal rates of TFP growth were very low – 0.4 between the 1900s and 1910s, 0.0 between the 1910s and the 1920s, and 0.2 between the 1920s and 1930s and there were spectacular jumps in sectoral rankings; for example, chemicals rose from 33$^{rd}$ to 4$^{th}$ and petroleum & coal products from 35$^{th}$ to 1$^{st}$ between the 1910s and the 1920s (Bakker et al., 2015). The insistence of Harberger (1998) that the process of TFP growth resembles 'mushrooms' rather than 'yeast' because it has 1001 different underpinnings seems appropriate.

In any case, as a robustness check, we allow for the possibility that the beginning-of-period industrial structure of cities is correlated with unobserved city-industry effects $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$. Similarly to Beaudry et al. (2012) or Bartik (1991), we then assume that, the *past* growth of national employment is not correlated with $\epsilon_{cs} = \epsilon_{cst+m} - \epsilon_{cst}$. This allows us to construct the predicted industrial structure of cities and use it as an instrument to estimate equation (10) with 2-stage least squares for samples beginning in 1890. Specifically, we predict the level of employment for industry *s* in city *c* in period *t+m* as

$$l_{cst+m}^{predict} = l_{cst}\left(\frac{l_{st+m}}{l_{st}}\right) \qquad (11)$$

where the employment in city *c* and industry *s* at time *t* is multiplied by the employment growth of that industry at the national level. This instrument breaks the direct link between city-level industrial structure in 1890 and the city-industry productivity growth rate in 1890-1920, by predicting the city's industrial structure with the past growth of national employment. We then use the predicted values of employment to calculate $MAR_{cs}$ and $JACOBS_{cs}$ as in equations (8) and (9), respectively.

Since we do not have data on cities in 1870, we cannot use this method to calculate predicted city-level employment in 1880. Therefore, in that case we use heteroskedasticity-based instruments as proposed by Lewbel (2012) and applied by, for example, Emran et al. (2013). The main idea is to use heteroskedasticity as identification when we do not have 'external' instruments. Consider the following system of two equations

$$Y_1 = X'\beta_1 + Y_2\gamma_1 + \varepsilon_1 \qquad (12)$$
$$Y_2 = X'\beta_2 + \varepsilon_2 \qquad (13)$$

Lewbel (2012) shows that, if it is assumed that $E(X\varepsilon_1) = 0$, $E(X\varepsilon_2) = 0$, and for some vector Z such that $cov(Z, \epsilon_1\epsilon_2) = 0$ and $cov(Z, \varepsilon_2^2) \neq 0$, 2-stage least squares will identify the above system using X and $[Z - E(Z)]\widehat{\varepsilon_2}$ as instruments. The requirement that $cov(Z, \varepsilon_2^2) \neq 0$ can be established with standard heteroskedasticity tests. An important thing for us is that vector Z can contain some or all elements of X (Lewbel 2012, p.70). Indeed, in his example of the estimation of an Engel curve, vector Z contains only elements of vector X. In our case, we estimate equation (10) with 2-stage least squares where the first-stage regressions are:

$$MAR_{cst} = \alpha_0 + \eta_0 X_{cst} + \theta_s + \lambda_c + \delta_{cs}^{MAR} \qquad (14)$$
$$JACOBS_{cst} = \alpha_0 + \eta_0 X_{cst} + \theta_s + \lambda_c + \delta_{cs}^{JACOBS} \qquad (15)$$

The estimated residuals are then used to construct the instruments $[Z - E(Z)]'\widehat{\delta_{cs}}$ where $\widehat{\delta_{cs}} = \left[\hat{\delta}_{cs}^{MAR}, \hat{\delta}_{cs}^{JACOBS}\right]$ and Z contains all variables on the right hand side of equation 14 or 15.

Table 10 provides the results of estimating equation (10) with the instruments offered by equations (11), (14) and (15). We see that the instruments pass the weak instrument test and that the estimated relationships between labor productivity and Marshallian and Jacobian externalities reported in Table 8 remain intact.

**Economic significance of results**

The estimates in Table 8 establish that greater specialization within cities tended to raise labor productivity growth in the United States in the early 20th century. On the other hand, greater diversity of manufacturing production had different effects depending on its extent, and tended to raise productivity in large cities but to reduce it in smaller cities. We have also seen that in the decades at the end of the 19th and beginning of the 20th century American cities were generally becoming more specialized and less diversified. Our regressions imply that the overall effect of these changes in composition of economic activity within cities tended significantly to increase productivity.

Table 11 reports estimates of the impact of changes in *MAR* and *JACOBS* scores for labor productivity in American cities using the estimated equation for 1890-1920 reported in Table 8, column 3. The result is that increased agglomeration externalities are estimated to have raised labor productivity in this sample of cities by 25.84 per cent in 1920 compared with 1890. Almost all of this came from increased specialization with the contribution of Δ*MAR* accounting for a 23.15 per cent addition to productivity. As for Jacobian externalities, we see a mixed picture. In smaller cities falls in JACOBS were positive for labor productivity but this was offset by the adverse effects of lower diversity in big cities. In any case, of course, productivity was not rising in this period as a result of increases in the diversity of production in American cities.

In Table 12 we show some of the implications of *MAR* and *JACOBS* remaining at their 1890 levels such that the increased productivity resulting from greater agglomeration externalities was not realized. The effect is sizeable relative to productivity advance in manufacturing at this time. If specialization and diversity of production in American cities were still at the 1890 levels, then labor productivity growth between 1890 and 1920 in manufacturing would have been reduced from 0.85 per cent per year (Kendrick, 1961) to 0.32 per cent per year and the ratio of labor productivity in 1920 in the United States relative to the United Kingdom would have fallen from 2.228 (Broadberry, 1997) to 1.916 – a little below where it had been in 1890.

To highlight further the economic significance of agglomeration externalities, we use counterfactual analysis in the spirit of Redding and Venables (2004) or Klein and Crafts (2012). As pointed out above, in the decades at turn of the 20th century, cities were becoming more specialized and less diversified; Detroit is a well-known example. However, Detroit was still quite a diversified city in 1900 before turning into the center of US automobile industry a couple of decades later. On the other hand, Akron, Ohio was already highly specialized by 1900 as it had become the center of US rubber industry and 29.6% of all workers in the city worked in rubber or tire production.[10] Indeed, the specialization index of Akron, Ohio was at the 80th while that of Detroit was at the 15th percentile. As a counterfactual analysis, we ask by how much would the productivity of Detroit have changed between 1900 and 1920 if it already had the degree of specialization as Akron, Ohio in 1900. Using the estimated coefficients for the period 1900-1920 with Akron, Ohio having a 60% higher specialization index than Detroit, the labor productivity of Detroit would have increased by 3.12% per annum.

## VI. Discussion

The literature offers little guide to how productivity may vary with the degree of industrial diversity in a city. The only exception is a conjecture offered by Fujita and Thisse (2013, p. 136) that small and medium size cities benefit mostly from intra-sectoral agglomeration economies while large cities benefit from both intra- and inter-sectoral agglomeration economies. The implication is that in small and medium-size cities productivity advantages should result from

[10] Calculated from 1900 US Census of Manufactures.

Marshallian rather than Jacobian externalities while in large cities both types of externalities should have positive effects. This conjecture is consistent with the theoretical analysis of Duranton and Puga (2001, Lemma 4) which proposes that diversified cities are a more costly place to produce than specialized cities. To see why, compare an increase in employment in the iron and steel industry in a city specialized in iron and steel production with a similar increase in a town where the iron and steel industry is one of many industries. The additional employment has both a cost-reducing and a cost-increasing effect. The former comes from the increased size of the industry, the latter from worsening congestion and higher labor costs. The magnitude of those effects differs in the two cities. In the specialized city, the employment increase just has own-industry effects of both types while in the diversified city it has a cost-reducing effect for the iron and steel industry but is cost-increasing for all other industries.

Our empirical results are consistent with these theoretical arguments. They imply that there is a city-size threshold below which the costs of diversification are larger than the benefits, and above which the benefits begin to exceed the costs. In other words, our results show that while Marshallian externalities always increase productivity whatever the city size, Jacobian externalities have a productivity-enhancing effect only for cities above a certain threshold below which their effect is productivity reducing. This finding about Jacobian externalities adds to a relatively small but growing literature about the heterogeneous effects of agglomeration externalities.[11] So far, this has identified heterogeneous agglomeration effects across industries (Brulhart and Mattys, 2008), workers and firms (Abel et al., 2012; Lee et al., 2010; Combes et al., 2012), or demographic groups (Ananat et al., 2013), but a non-linear effect of Jacobian externalities on labor productivity across cities of various size has not previously been found.

These results also have implications for the well-known analysis of American economic growth in Lucas (1988) which emphasized the importance of Jacobian externalities as an engine of growth. To some extent, our findings provide support for this hypothesis during the period of the second industrial revolution. The key point that stands out is that these externalities from diversity were not generated by urbanization generally but accrued in the really large cities of the United States. If Lucas's vision is correct, the basis for it would be the economic activity of Chicago, New York and Philadelphia rather than Akron, Ohio or Hartford, Connecticut. On the whole our evidence gives stronger support to the role of Marshallian externalities based on specialization as the key contribution of cities to American economic growth.

We find that increases in specialization in cities allowed a significant in labor productivity gain in manufacturing in the period 1890-1920. What led cities to become more specialized? This is an important topic for future quantitative research but the traditional history literature provides a string hypothesis. Pred (1966) stressed the role of falling railroad rates which he argued increased the practicability of serving national demands from a limited number of cities. The extension and enlargement of a firm's market increased the feasibility of agglomeration and large scale production at this time. Similarly, Chandler (1977) saw the decades before World War I as inaugurating an ear of mass production combined with mass distribution which saw the rise of national brands like Pillsbury Flour, Heinz, Procter and Gamble etc. based in cities such as Minneapolis, Pittsburgh and Cincinnati.

[11] See Combes and Gobillon (2014) for a useful survey.

If this conjecture is correct, it has important ramifications for a famous debate in economic history, namely, the contribution of railroads to American economic growth. Fogel (1964) estimated this using the technique of social savings which essential used the transport benefits of railroads resulting from a lower price of transportation as a measure of their overall economic benefits. This is only valid under quite strong assumptions which include that the rest of the economy is perfectly competitive with constant returns to scale (Jara-Diaz, 1986). Agglomeration economies in the transport-using sector (wider economic benefits) are not captured by Fogel's technique. Our research suggests that the contribution of railroads to American economic growth may have been larger than Fogel supposed.[12]

## VII. Conclusions

Our analysis of labor productivity growth in early 20th-century U.S. cities has produced some interesting new results. The main conclusions are the following.

First, we find that on average more specialized cities enjoyed higher labor productivity growth in manufacturing. This suggests that Marshallian externalities were an important aspect of urbanization during the second industrial revolution. An important task for future research will be to explore the sources of these agglomeration economies.

Second, the relationship between diversity of industrial structure and manufacturing labor productivity growth appears to have been non-linear. It seems likely that Jacobian externalities were only realized in large cities such as Chicago, Philadelphia, and New York. In smaller cities, increased diversity tended to reduce productivity growth.

Third, the degree of industrial specialization in American cities increased considerably in the early 20th century. This had a significant impact on labor productivity. We estimate that if specialization had remained at the 1890 level the 1920 level of manufacturing labor productivity in the United States would have been about 25 per cent lower.

[12] Fogel (1964) provided what he claimed was an upper bound estimate of social savings (user benefits) which he obtained by assuming a vertical demand curve for the volume of goods to be transported. If, as we suggest, there were important externalities in the transport-using sector which were a result of the development of railroad technology, then Fogel's estimate is not necessarily an upper bound.

## References

Abdel-Rahman, H. M. (1996), "When Do Cities Specialize in Production?", *Regional Science and Urban Economics*, 26, 1-22.

Abel, J.R, Ishita, D., and Todd, G. (2012), "Productivity and the Density of Human Capital", *Journal of Regional Science*, 52, 562-586.

Ananat, E., Shine, F., and Ross, S.L. (2013), "Race-Specific Agglomeration Economies: Social Distance and the Black-White Wage Gap", NBER Working Paper 18933.

Bakker, G., Crafts, N. and Woltjer, P. (2015), "A Vision of the Growth Process in a Technologically Progressive Economy: the United States, 1899-1941", CEPR Working Paper.

Bartik, T. (1991), *Who Benefits from State and Local Economic Development Policies?* Kalmzoo, MI: Upjohn Institute.

Beaudry, P., Green, D., and Sand, B. (2012), "Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory", *Econometrica*, 80, 1063-1104.

Bostic, R. W., Gans, J. S. and Stern, S. (1997), "Urban Productivity and Factor Growth in the Late Nineteenth Century", *Journal of Urban Economics*, 41, 38-55.

Broadberry, S. N. (1997), *The Productivity Race*. Cambridge: Cambridge University Press.

Brülhart, M., and Mathys, N.A. (2008), "Sectoral Agglomeration Economies in a Panel of European Regions", *Regional Science and Urban Economics*, 38, 348-362.

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2011), "Robust Inference with Multi-Way Clustering", *Journal of Business and Economic Statistics*, 29, 238-249.

Chandler, A. D. (1977), *The Visible Hand*. Cambridge, Mass.: Belknap Press.

Cingano, F. and Schivardi, F. (2004), "Identifying the Sources of Local Productivity Growth", *Journal of the European Economic Association*, 2, 720-742.

Combes, P-P. (2000), "Economic Structure and Local Growth: France, 1981-1993", *Journal of Urban Economics*, 47, 329-355.

Combes, P-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012), "The Productivity Advantages of Large Markets: Distinguishing Agglomeration from Firm Selection", *Econometrica*, 80, 2543-2594.

Combes, P-P. and Gobillon, L. (2014), "The Empirics of Agglomeration Economies", CEPR Discussion Paper No. 10174.

Combes, P-P., Mayer, T. and Thisse, J-F. (2008), *Economic Geography: the Integration of Regions and Nations*. Princeton: Princeton University Press.

David, P. A. (1991), "Computer and Dynamo: the Modern Productivity Paradox in a Not-Too-Distant Mirror", in OECD, *Technology and Productivity: the Challenge for Economic Policy*. Paris: OECD, 315-347.

Duranton, G. (2014), Growing through Cities in Developing Countries", World Bank Policy Research Working Paper No. 6818.

Duranton, G. and Puga, D. (2001), "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products", *American Economic Review*, 91, 1454-1477.

Duranton, G. and Puga, D. (2004), "Micro-Foundations of Urban Agglomeration Economies", in J. V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, 2063-2117.

Duranton, G. and Puga, D. (2013), "The Growth of Cities", CEPR Discussion Paper No. 9590.

Emran, M.S., and Hou, Z. (2013), "Access to Markets and Rural Poverty: Evidence from Household Consumption in China", *Review of Economics and Statistics*, 95, 682-697.

Fogel, R. W. (1964), *Railroads and American Economic Growth*. Baltimore: Johns Hopkins Press.

Fujita, M. and Thisse, J-F. (2013), *Economics of Agglomeration, 2nd edition*. Cambridge: Cambridge University Press.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. and Shleifer, A. (1992), "Growth in Cities", *Journal of Political Economy*, 100, 1126-1152.

Harberger, A. C. (1998), "A Vision of the Growth Process", *American Economic Review*, 88, 1-32.

Harris, R. and Moffat, J. (2012), "Is Productivity Higher in British Cities?", *Journal of Regional Science*, 52, 762-786.

Henderson, J. V. (1974), "The Sizes and Types of Cities", *American Economic Review*, 64, 640-656.

Henderson, J. V. (2003), "Marshall's Scale Economies", *Journal of Urban Economics*, 53, 1-28.

Henderson, J. V., Kuncor, A. and Turner, M. (1995), "Industrial Development in Cities", *Journal of Political Economy*, 103, 1067-1090.

Jacobs, J. (1969), *The Economy of Cities*. New York: Random House.

Jara-Diaz, S.R. (1986), "On the Relation between Users' Benefits and the Economic Effects of Transportation Activities", *Journal of Regional Science*, 26, 379-391.

Kendrick, J. W. (1961), *Productivity Trends in the United States*. Princeton, NJ.: Princeton University Press.

Kim, S. (2000), "Urban Development in the United States, 1690-1990", *Southern Economic Journal*, 66, 855-880.

Kim, S. (2006), "Division of Labor and the Rise of Cities: Evidence from US Industrialization, 1850-1880", *Journal of Economic Geography*, 6, 469-491.

Kim, S. and Margo, R. A. (2004), "Historical Perspectives on U. S. Economic Geography", in J. V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, 2981-3019.

Klein, A. and Crafts, N. (2012), "Making Sense of the Manufacturing Belt: Determinants of US Industrial Location, 1880-1920", *Journal of Economic Geography* (2012), 12, 775-807.

Lee, B.S., Jang, S., and Hong, S.H. (2010), "Marshall Scale Economies and Jacobs Externalities in Korea: The Role of Age, Size, and the Legal Form of Establishments", *Urban Studies*, 47, 3131-3156.

Lewbel, A. (2012), "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models", *Journal of Business and Economic Statistics*, 30, 67-80.

Lucas, R. E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22, 3-42.

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

Pred, A. R. (1966), *The Spatial Dynamics of U. S. Urban-Industrial Growth, 1800-1914*. Cambridge, Mass.: MIT Press.

Redding, S. and Venables, A. J. (2004), "Economic Geography and International Inequality", *Journal of International Economics*, 62, 53–82.

Robinson, P. M. (1988), "Root-N-Consistent Semiparametric Regression" *Econometrica* 56, 931-954.

Rosenthal, S. S. and Strange,W. C. (2004), "Evidence on the Nature and Sources of Agglomeration Economies", in J. V. Henderson and J-V. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, 2119-2171.

## Appendix

The results in Table 8 show a quadratic relationship between average annual labor productivity and Jacobian externalities. To explore the robustness of that relationship, we dispose of the quadratic specification of Jacobian externalities, make no assumption about that relationship and estimate a semi-parametric specification of equation (10)

$$ln\frac{vapw_{cst+m}}{vapw_{cst}} = \alpha + \beta_1 MAR_{cst} + f(JACOBS_{cst}) + \beta_2 X_{cst} + \theta_s + \lambda_c + \epsilon_{cs} \quad (11).$$

We estimate equation (11) with Robinson's (1988) semi-parametric estimator which uses a double-residual methodology. We present the results of that non-parametric function only in Figure 6 since the results for the specialization index and the other controls are qualitatively unchanged. The results reveal interesting patterns. The non-parametric estimates show broadly a curvilinear pattern for the periods 1890-1920 and 1900-1930 though the pattern in 1900-1930 exhibits two local u-shaped patterns for relatively less diversified towns. This, however, is driven by small number of observations in the left part of the distribution with towns having a diversity index around 1.8 and 2.3, respectively, which, with different non-parametric bandwidth might yield a smoother curvilinear pattern. The patterns in those time periods are confirmed by the parametric estimates. The period 1880-1920 also shows a curvilinear pattern, though not as pronounced as in 1890-1920 or 1900-1930. The period 1880-1930 shows no relationship between industrial diversity and average annual labor productivity growth, similar to the parametric estimates, while 1900-1920 exhibits a declining pattern which peters out for highly diversified cities. All this is consistent with the parametric estimates. The period 1890-1930 is the only period exhibiting a clear U-shaped relationship for which statistical significance is not confirmed by the parametric regression.[13] Overall, the graphs show that in four out of six cases, Jacobian externalities exhibit a non-linear U-shaped pattern in which the average annual labor productivity begins to increase for rather diversified cities with industrial variety indices between 2 and 2.5, similar to those reported in Table 10.

[13] The graphs of nonparametric estimates in Figure 6 were trimmed from below for the period 1890-1930 and 1900-1930 respectively due to a handful of observations being clear outliers. The graphs including the outliers are available from the authors upon request.

**Table 1: Number of Cities by US Regions as Reported in the US Census of Manufactures 1880-1930.**

| Region/Number of Cities | 1880 | 1890 | 1900 | 1920 | 1930 |
|---|---|---|---|---|---|
| East North Central | 18 | 33 | 44 | 26 | 8 |
| East South Central | 6 | 11 | 11 | 8 | 1 |
| Middle Atlantic | 26 | 40 | 53 | 35 | 15 |
| Mountain | 2 | 3 | 4 | 1 | 1 |
| New England | 21 | 32 | 40 | 22 | 10 |
| Pacific | 3 | 7 | 9 | 11 | 8 |
| South Atlantic | 10 | 12 | 16 | 14 | 3 |
| West North Central | 8 | 18 | 23 | 14 | 6 |
| West South Central | 3 | 7 | 9 | 9 | 7 |
| DC | 1 | 1 | 0 | 0 | 0 |
| *US* | *98* | *164* | *209* | *140* | *59* |

*Sources*: US Census of Manufactures 1880, 1890, 1900, 1920, 1930

**Table 2: Number of Matched Cities 1880-1930.**

| Number of Cities/Time Period | 1880-1920 | 1880-1930 | 1890-1920 | 1890-1930 | 1900-1920 | 1900-1930 |
|---|---|---|---|---|---|---|
| Matched cities | 81 | 41 | 118 | 53 | 127 | 54 |
| Cities entering sample in 1890 and 1900 | | | 37 | 12 | 9 | 1 |
| Cities exiting the sample in 1930 | | 40 | | 65 | | 73 |

*Sources*: US Census of Manufactures 1880, 1890, 1900, 1920, 1930

**Table 3: Average share of employment in selected industries in US Cities 1880, 1930 (%).**

| SIC 3 | Industry SIC 3 | 1880 | 1930 |
|---|---|---|---|
| | | *Traditional Industries* | |
| 203 | Preserved Fruits & Vegetables | 7.94 | 2.45 |
| 208 | Beverages | 2.63 | 0.49 |
| 223 | Broadwoven Fabric Mills, Wool | 12.68 | 5.32 |
| 225 | Knitting Mills | 1.78 | 0.43 |
| 242 | Sawmills & Planing Mills | 6.75 | 1.79 |
| 273 | Books | 5.28 | 2.55 |
| 313 | Footwear Cut Stock | 4.41 | 2.57 |
| 319 | Leather Goods, nec | 1.23 | 0.19 |
| 325 | Structural Clay Products | 3.55 | 1.24 |
| 332 | Iron Foundries | 13.81 | 6.27 |
| | | *Industries of 'Second Industrial Revolution'* | |
| 287 | Agricultural Chemicals | 0.65 | 1.10 |
| 291 | Petroleum Refining | 0.00 | 20.02 |
| 331 | Blast Furnace & Basic Steel Products | 0.53 | 6.00 |
| 345 | Screw Machine Products, Bolts, Etc. | 0.62 | 2.13 |
| 355 | Special Industry Machinery | 0.63 | 2.63 |
| 356 | General Industry Machinery | 0.18 | 0.26 |
| 362 | Electrical Industrial Apparatus | 0.32 | 3.68 |
| 364 | Electric Lighting & Wiring Equipment | 0.27 | 0.48 |
| 371 | Motor Vehicles & Equipment | 1.41 | 8.77 |
| 382 | Measuring & Controlling Devices | 0.18 | 0.89 |

*Sources*: US Census of Manufactures 1880, 1930.

*Note*: The employment shares are calculate relative to city total employment.

**Table 4: Average Annual Growth of Value-Added per Worker of SIC 3 Industries in US Cities: 1880-1920.**

| US State | US City | SIC 3 | Industry Name SIC 3 | Growth Rate 1880-1920 |
|---|---|---|---|---|
| | | | *Fastest Growing Industries* | |
| NEW YORK | BUFFALO | 242 | Sawmills & Planing Mills | 0.077 |
| PENNSYLVANIA | PHILADELPHIA | 339 | Misc Primary Metal Industries | 0.061 |
| KENTUCKY | LOUISVILLE | 289 | Miscellaneous Chemical Products | 0.056 |
| GEORGIA | ATLANTA | 399 | Miscellaneous Manufacturers | 0.055 |
| OHIO | CLEVELAND | 342 | Cutlery, Handtools & Hardware | 0.054 |
| NEW YORK | BUFFALO | 331 | Blast Furnace & Basic Steel Products | 0.053 |
| NEW JERSEY | NEWARK | 284 | Soap, Cleaners & Toilet Goods | 0.053 |
| NEW JERSEY | Hoboken | 332 | Iron & Steel Foundries | 0.051 |
| NEW YORK | BUFFALO | 283 | Drugs | 0.051 |
| PENNSYLVANIA | SCRANTON | 332 | Iron & Steel Foundries | 0.050 |
| MICHIGAN | DETROIT | 399 | Miscellaneous Manufacturers | 0.050 |
| NEW YORK | UTICA | 344 | Fabricated Structural Metal Products | 0.050 |
| PENNSYLVANIA | PHILADELPHIA | 286 | Industrial Organic Chemicals | 0.049 |
| MICHIGAN | DETROIT | 382 | Measuring & Controlling Devices | 0.048 |
| | | | *Slowest Growing Industries* | |
| NEW YORK | NEW YORK CITY | 207 | Fats & Oils | -0.021 |
| MASSACHUSETTS | NEW BEDFORD | 201 | Meat Products | -0.023 |
| MASSACHUSETTS | WORCESTER | 319 | Leather Goods, nec | -0.023 |
| RHODE ISLAND | Providence | 209 | Misc Food & Kindred Products | -0.023 |
| INDIANA | INDIANAPOLIS | 201 | Meat Products | -0.023 |
| NEW YORK | ROCHESTER | 209 | Misc Food & Kindred Products | -0.023 |
| NEW JERSEY | NEWARK | 201 | Meat Products | -0.023 |
| NEW YORK | BUFFALO | 275 | Commercial Printing | -0.023 |
| MISSOURI | Kansas city | 242 | Sawmills & Planing Mills | -0.026 |
| MASSACHUSETTS | Lawrence | 251 | Household Furniture | -0.026 |
| CALIFORNIA | SAN FRANCISCO | 289 | Miscellaneous Chemical Products | -0.027 |
| MINNESOTA | Saint Paul | 278 | Blankbooks & Bookbinding | -0.029 |
| OHIO | CINCINNATI | 259 | Miscellaneous Furniture & Fixtures | -0.029 |
| CALIFORNIA | SAN FRANCISCO | 201 | Meat Products | -0.029 |

*Sources*: US Census of Manufactures 1880, 1920.

**Table 5: Average Annual Growth Rate of Value-Added per Worker of SIC 2 Industries in US Cities, 1880-1930.**

| SIC 2 | Industry SIC2 | Growth Rate of Value-Added per Worker | | | |
|---|---|---|---|---|---|
| | | 1880-1920 | 1880-1930 | 1890-1920 | 1890-1930 |
| 20 | Food and kindred Product | 0.001 | 0.019 | -0.015 | 0.010 |
| 21 | Tobacco and tobacco product | 0.020 | 0.017 | -0.003 | 0.008 |
| 22 | Textile mill product | 0.008 | 0.021 | -0.002 | 0.020 |
| 23 | Apparel and related products | 0.020 | 0.025 | 0.003 | 0.015 |
| 24 | Lumber and wood products | -0.004 | 0.021 | -0.015 | 0.014 |
| 25 | Furniture and fixtures | -0.004 | 0.020 | -0.015 | 0.015 |
| 26 | Paper and allied products | -0.003 | 0.027 | -0.011 | 0.022 |
| 27 | Printing and publishing | -0.006 | 0.025 | -0.008 | 0.026 |
| 28 | Chemicals and allied products | 0.022 | 0.026 | 0.002 | 0.015 |
| 29 | Petroleum and coal products | 0.014 | 0.027 | -0.006 | 0.021 |
| 30 | Rubber and plastic products | 0.016 | 0.018 | 0.000 | 0.021 |
| 31 | Leather and leather products | 0.008 | 0.019 | -0.015 | 0.012 |
| 32 | Stone, clay, and glass products | 0.010 | 0.023 | 0.000 | 0.017 |
| 33 | Primary metal products | 0.033 | 0.023 | 0.022 | 0.015 |
| 34 | Fabricated metal products | 0.029 | 0.021 | 0.016 | 0.013 |
| 35 | Machinery | 0.026 | 0.020 | 0.017 | 0.007 |
| 36 | Electrical equipment | 0.026 | 0.025 | 0.017 | 0.020 |
| 37 | Transportation equipment | 0.027 | 0.016 | 0.021 | 0.008 |
| 38 | Instruments and related products | 0.017 | 0.021 | -0.002 | 0.007 |
| 39 | Miscellaneous manufacturing including ordnance | 0.031 | 0.023 | 0.019 | 0.014 |

*Sources*: US Census of Manufactures 1880, 1890, 1900, 1920, 1930

*Note*: Growth rates calculated as the average of SIC 3 industries.

**Table 6: Top five and bottom five industries by specialization and variety index.**

| 1880 | 1900 | 1930 |
|---|---|---|
| | **Panel A:** | |
| | *SIC 3 industries with the highest specialization index* | |
| Miscellaneous Manufacturers | Miscellaneous Manufacturers | Miscellaneous Manufacturers |
| Broadwoven Fabric Mills, Cotton | Broadwoven Fabric Mills, Cotton | Petroleum Refining |
| Iron & Steel Foundries | Petroleum Refining | Motor Vehicles & Equipment |
| Broadwoven Fabric Mills, Wool | Broadwoven Fabric Mills, Wool | Iron & Steel Foundries |
| Chewing & Smoking Tobacco | Iron & Steel Foundries | Blast Furnace & Basic Steel Products |
| | *SIC 3 industries with the lowest specialization index* | |
| Greeting Cards | Handbags & Personal Leather Goods | Miscellaneous Primary Metal Industries |
| Industrial Machinery, nec | Secondary Nonferrous Metals | Industrial Machinery, nec |
| Nonferrous Foundries (Castings) | Industrial Organic Chemicals | Refrigeration & Service Industry |
| Photographic Equipment & Supplies | General Industry Machinery | Watches, Clocks, Watchcases & Parts |
| Plastics Materials & Synthetics | Household Audio & Video Equipment | Miscellaneous Transportation Equipment |
| | **Panel B:** | |
| | *SIC 3 industries with the highest variety index* | |
| Plastics Materials & Synthetics | Handbags & Personal Leather Goods | Industrial Machinery, nec |
| Fabricated Rubber Products, nec | Office Furniture | Miscellaneous Primary Metal Industries |
| Toys & Sporting Goods | Photographic Equipment & Supplies | Miscellaneous Textile Goods |
| Leather Gloves & Mittens | Household Appliances | Knitting Mills |
| Special Industry Machinery | Nonferrous Foundries (Castings) | Carpets & Rugs |
| | *SIC 3 industries with the lowest variety index* | |
| Metal Forgings & Stampings | Leather Goods, nec | Aircraft & Parts |
| Broadwoven Fabric Mills, Wool | Newspapers | Special Industry Machinery |
| Iron & Steel Foundries | Metal Forgings & Stampings | Books |
| Paper Mills | Bakery Products | Bakery Products |
| Tobacco Stemming & Redrying | Men's & Boys' Furnishings | Iron & Steel Foundries |

*Sources*: US Census of Manufactures 1880, 1900, 1930.

*Note*: Specialization and variety indices are calculated as the average across all cities in the data set.

**Table 7: Agglomeration and Productivity Growth: OLS Regressions**

| | 1880-1920 | 1880-1930 | 1890-1920 | 1890-1930 | 1900-1920 | 1900-1930 |
|---|---|---|---|---|---|---|
| *JACOBS* | -0.00315** | 0.00001 | -0.00798*** | 0.00215 | -0.02139*** | -0.00499 |
| | [0.001] | [0.001] | [0.002] | [0.003] | [0.005] | [0.005] |
| *MAR* | 0.01440*** | 0.00881* | 0.02626*** | 0.01149 | 0.05343*** | 0.02723*** |
| | [0.005] | [0.005] | [0.007] | [0.009] | [0.012] | [0.007] |
| Initial Value-Added per Worker | -0.00002*** | -0.00001*** | -0.00001*** | -0.00001*** | -0.00001* | -0.00000** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Initial State Manufacturing Employment (% US) | 0 | -0.00007 | 0.00001 | -0.00001 | -0.00007 | 0.00003 |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Constant | 0.03385*** | 0.03537*** | 0.01528*** | 0.01356*** | 0.03614*** | 0.02197*** |
| | [0.003] | [0.004] | [0.006] | [0.004] | [0.008] | [0.007] |
| City Dummies | YES | YES | YES | YES | YES | YES |
| Industry Dummies | YES | YES | YES | YES | YES | YES |
| *Joint significance* | | | | | | |
| *JACOBS* & *MAR* ≠ 0 | 8.93** | 3.38 | 15.99*** | 4.16^ | 19.82*** | 17.41*** |
| $R^2$ | 0.609 | 0.635 | 0.525 | 0.677 | 0.432 | 0.508 |
| N | 1,254 | 727 | 1,954 | 1,046 | 2,192 | 1,157 |

*Sources*: US Census of Manufactures 1880, 1890, 1900, 1920, 1930.

*Notes*: ^ significant at 12%, * significant at 10%; ** significant at 5%; *** significant at 1%

standard errors are clustered at city and industry level

**Table 8: Agglomeration and Productivity Growth: OLS Regression with Nonlinearities.**

| | 1880-1920 | 1880-1930 | 1890-1920 | 1890-1930 | 1900-1920 | 1900-1930 |
|---|---|---|---|---|---|---|
| *JACOBS* | -0.01744** | -0.01266 | -0.02331*** | -0.00883 | -0.02541*** | -0.03045*** |
| | [0.008] | [0.013] | [0.007] | [0.009] | [0.007] | [0.009] |
| *$JACOBS^2$* | 0.00444* | 0.00356 | 0.00453** | 0.00297 | 0.00127 | 0.00689** |
| | [0.003] | [0.004] | [0.002] | [0.002] | [0.002] | [0.003] |
| *MAR* | 0.01633*** | 0.01021* | 0.03088*** | 0.0129^ | 0.05421*** | 0.02967*** |
| | [0.005] | [0.006] | [0.008] | [0.009] | [0.012] | [0.008] |
| Initial Value-Added/Worker | -0.00002*** | -0.00001*** | -0.00001*** | -0.00001*** | -0.00001* | -0.00000** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Initial State Manufacturing Employment (% US) | 0 | -0.00007 | 0 | -0.00001 | -0.00007 | 0.00001 |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Constant | 0.04503*** | 0.04405*** | 0.02685*** | 0.02240*** | 0.03910*** | 0.04297*** |
| | [0.007] | [0.010] | [0.008] | [0.007] | [0.009] | [0.008] |
| *Joint significance* | | | | | | |
| *JACOBS* & *$JACOBS^2$* $\neq 0$ | 8.92** | 1.09 | 23.54*** | 2.43 | 21.78*** | 10.77*** |
| *JACOBS*, *$JACOBS^2$* & *MAR* $\neq 0$ | | 5.51^ | | 6.27* | | |
| City Dummies | YES | YES | YES | YES | YES | YES |
| Industry Dummies | YES | YES | YES | YES | YES | YES |
| $R^2$ | 0.611 | 0.636 | 0.526 | 0.678 | 0.432 | 0.51 |
| N | 1,254 | 727 | 1954 | 1,046 | 2,192 | 1,157 |

*Sources*: US Census of Manufactures 1880, 1890, 1900, 1920, 1930.

*Notes*: ^ significant at 13%, * significant at 10%; ** significant at 5%; *** significant at 1%

standard errors are clustered at city and industry level

**Table 9: Agglomeration and Productivity Growth: OLS Regression. Robustness Check**

| | 1890-1920 | 1890-1930 | 1900-1920 | 1900-1930 | 1900-1920 | 1900-1930 |
|---|---|---|---|---|---|---|
| | Sample of Cities in 1880 | | Sample of Cities in 1880 | | Sample of Cities in 1890 | |
| | *Basic specification* | | | | | |
| *JACOBS* | -0.00943*** | 0.00255 | -0.01575*** | -0.00302 | -0.01505*** | -0.00193 |
| | [0.002] | [0.004] | [0.005] | [0.004] | [0.005] | [0.004] |
| *MAR* | 0.03406*** | 0.01756** | 0.04331*** | 0.01966*** | 0.03842*** | 0.01372** |
| | [0.008] | [0.009] | [0.013] | [0.006] | [0.011] | [0.006] |
| Value-added/Worker | -0.00001*** | -0.00001*** | -0.00002*** | -0.00001*** | -0.00002*** | -0.00001*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| State Manf. Empt. (% US) | 0 | -0.00003 | -0.00003 | 0.00001 | -0.00004 | 0.00003 |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Constant | 0.02552*** | 0.02003*** | 0.05971*** | 0.04893*** | 0.05131*** | 0.03944*** |
| | [0.004] | [0.005] | [0.007] | [0.008] | [0.006] | [0.005] |
| *Joint significance* | | | | | | |
| *JACOBS* & *MAR* ≠ 0 | 24.89*** | 5.63* | 11.97** | 9.57** | 12.12*** | 6.45** |
| City Dummies | YES | YES | YES | YES | YES | YES |
| Industry Dummies | YES | YES | YES | YES | YES | YES |
| $R^2$ | 0.528 | 0.663 | 0.56 | 0.642 | 0.553 | 0.634 |
| N | 1,581 | 876 | 1,693 | 935 | 2,138 | 1,148 |
| | *Nonlinearities* | | | | | |
| *JACOBS* | -0.02282*** | -0.00263 | -0.01462* | -0.02660** | -0.01977*** | -0.02642*** |
| | [0.009] | [0.010] | [0.008] | [0.013] | [0.007] | [0.008] |
| $JACOBS^2$ | 0.00394 | 0.0014 | -0.00034 | 0.00616** | 0.00144 | 0.00663*** |
| | [0.003] | [0.003] | [0.002] | [0.003] | [0.002] | [0.002] |
| MAR | 0.03651*** | 0.01756** | 0.04319*** | 0.01823*** | 0.03946*** | 0.01617** |
| | [0.008] | [0.009] | [0.013] | [0.006] | [0.011] | [0.006] |
| Value-added/Worker | -0.00001*** | -0.00001*** | -0.00002*** | -0.00001*** | -0.00002*** | -0.00001*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| State Manf. Empt. (% US) | | | -0.00003 | 0.00001 | -0.00004 | 0.00002 |
| | | | [0.000] | [0.000] | [0.000] | [0.000] |
| Constant | 0.03574*** | 0.02374*** | 0.05878*** | 0.07085*** | 0.05486*** | 0.05964*** |
| | [0.008] | [0.007] | [0.009] | [0.015] | [0.007] | [0.007] |
| *Joint significance* | | | | | | |
| *JACOBS* & $JACOBS^2$ ≠ 0 | 33.25*** | 0.73 | 11.75*** | 4.21^ | 10.89** | 9.85* |
| *JACOBS*, $JACOBS^2$ & *MAR* ≠ 0 | | 6.27* | | | | |
| City Dummies | YES | YES | YES | YES | YES | YES |
| Industry Dummies | YES | YES | YES | YES | YES | YES |
| $R^2$ | 0.529 | 0.663 | 0.56 | 0.643 | 0.553 | 0.635 |
| N | 1,581 | 876 | 1,693 | 935 | 2,138 | 1,148 |

*Sources*: US Census of Manufactures 1890, 1900, 1920, 1930.

*Notes*: ^ significant at 12%, * significant at 10%; ** significant at 5%; *** significant at 1%
standard errors are clustered at city and industry level

**Table 10: Agglomeration and Productivity Growth: 2SLS Instrumental Variable Regressions.**

| | 1880-1920 | 1880-1930 | 1890-1920 | 1890-1930 | 1900-1920 | 1900-1930 |
|---|---|---|---|---|---|---|
| *JACOBS* | -0.03920** | -0.04060*** | -0.04106** | -0.00165 | -0.02491** | -0.01747* |
| | [0.017] | [0.012] | [0.018] | [0.024] | [0.012] | [0.010] |
| *$JACOBS^2$* | 0.00932** | 0.00916*** | 0.0071^ | 0.0007 | 0.00239 | 0.00517** |
| | [0.004] | [0.003] | [0.005] | [0.006] | [0.003] | [0.003] |
| *MAR* | 0.03143*** | 0.03738*** | 0.04601*** | 0.01819* | 0.03910*** | 0.01679** |
| | [0.011] | [0.007] | [0.011] | [0.009] | [0.012] | [0.008] |
| Initial Value-Added/Worker | -0.00002*** | -0.00001*** | -0.00001*** | -0.00001*** | -0.00002*** | -0.00001*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Initial State Manufacturing Employment (% US) | -0.00003 | -0.00013** | -0.00002 | -0.00001 | -0.00003 | 0.00002 |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| City Dummies | YES | YES | YES | YES | YES | YES |
| Industry Dummies | YES | YES | YES | YES | YES | YES |
| *Weak Instrument Tests* | | | | | | |
| *JACOBS (F-stat from first-stage)* | 31.76 | 72.27 | 17.87 | 12.55 | 69.42 | 39.98 |
| *$JACOBS^2$ (F-stat from first-stage)* | 56.26 | 138.27 | 24.57 | 20.21 | 66.84 | 38.19 |
| *MAR (F-stat from first-stage)* | 35.23 | 52.52 | 55.97 | 104.95 | 122.85 | 78.14 |
| *Anderson-Rubin chi-square test* | 13.45 | 68.74 | 26.25 | 27.78 | 16.55 | 9.528 |
| *Anderson-Rubin chi-square test p-values* | 0.00375 | 0 | 8.45E-06 | 4.03E-06 | 0.000873 | 0.023 |
| $R^2$ | 0.412 | 0.352 | 0.295 | 0.273 | 0.357 | 0.387 |
| N | 1,254 | 727 | 1,190 | 695 | 1,840 | 1,003 |

*Sources*: US Census of Manufactures 1880, 1890, 1900, 1920, 1930

*Note*: ^ significant at 12%; * significant at 10%; ** significant at 5%; *** significant at 1%

Standard errors are clustered at city and industry level; in 1880-1920 and 1880-1930, 2SLS is heteroskedasticity-based estimator.

The instrument is $(Z-E(Z))\varepsilon_2$ where Z is share of SIC2 manufacturing employment in US state, and $\varepsilon_2$ is the estimated residuals from the regression of endogenous on exogenous variables.

**Table 11. The Impact of Changes in MAR and JACOBS on Labor Productivity Growth in U. S. Cities, 1890-1920.**

| | |
|---|---|
| ***MAR*** | |
| Estimated Coefficient | 0.031 |
| Δ*MAR (percent)* | 24.9 |
| Impact on Productivity Growth (% points p.a.) | 0.77 |
| ***JACOBS*** | |
| Estimated Coefficient (*JACOBS*) | -0.2331 |
| Estimated Coefficient ($JACOBS^2$) | 0.00453 |
| *10th Percentile* | |
| Marginal Effect | -0.014 |
| Δ*JACOBS (percent)* | -37.2 |
| Impact on Productivity Growth (% points p.a.) | 0.52 |
| *25th Percentile* | |
| Marginal Effect | -0.009 |
| Δ*JACOBS (percent)* | -34.1 |
| Impact on Productivity Growth (% points p.a.) | 0.31 |
| *50th Percentile* | |
| Marginal Effect | -0.004 |
| Δ*JACOBS (percent)* | -26 |
| Impact on Productivity Growth (% points p.a.) | 0.11 |
| *75th Percentile* | |
| Marginal Effect | 0.003 |
| Δ*JACOBS (percent)* | -21.9 |
| Impact on Productivity Growth (% points p.a.) | -0.06 |
| *90th Percentile* | |
| Marginal Effect | 0.004 |
| Δ*JACOBS (percent)* | -11.3 |
| Impact on Productivity Growth (% points p.a.) | -0.04 |
| *95th Percentile* | |
| Marginal Effect | 0.004 |
| Δ*JACOBS (percent)* | -8.9 |
| Impact on Productivity Growth (% points p.a.) | -0.04 |
| *99th Percentile* | |
| Marginal Effect | 0.006 |
| *Δ JACOBS (percent)* | -11.1 |
| Impact on Productivity Growth (% points p.a.) | -0.1 |
| ***Cumulative Impact on Labor Productivity, 1890-1920 (%)*** | |
| *ΔMAR* | 23.15 |
| *ΔJACOBS* | 2.69 |
| Total | 25.84 |

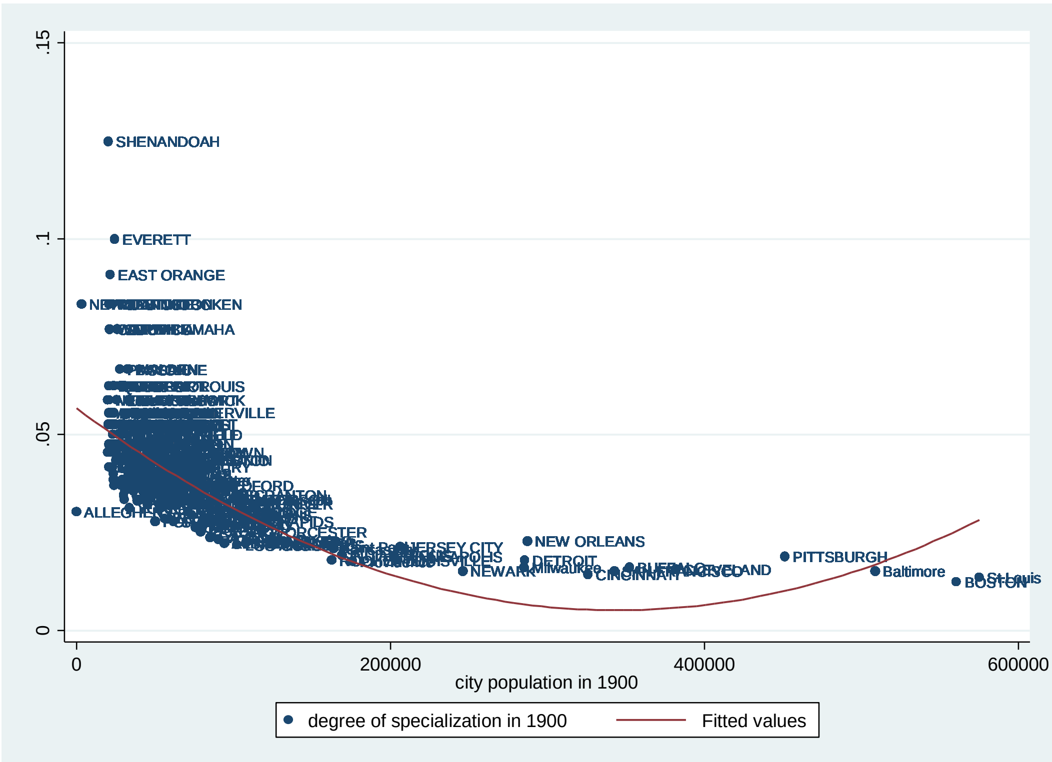
*Notes*: cumulative impact of Δ*JACOBS* is based on a weighted average of the estimates for each of the 7 percentile points.
The weights are the number of workers in a particular percentile group as a share of the number of workers in all cities.
*Source*: derived using the estimated equation in Table 8, column 3.

**Table 12. Counterfactual U.S. Manufacturing Labor Productivity in 1920.**

| | ***Actual*** | ***1890 MAR*** | ***1890 MAR and JACOBS*** |
|---|---|---|---|
| Output in Cities ($bn.) | 16.272 | 13.213 | 12.936 |
| Output/Worker in Cities ($) | 3000.2 | 2436.2 | 2385.2 |
| Total Output ($ bn.) | 23.842 | 20.783 | 20.506 |
| Total Output/Worker | 2816.5 | 2455.2 | 2422.5 |
| Labor Productivity Growth, 1890-1920 (% p.a.) | 0.85 | 0.36 | 0.32 |
| US/UK Labor Productivity Level | 2.228 | 1.942 | 1.916 |

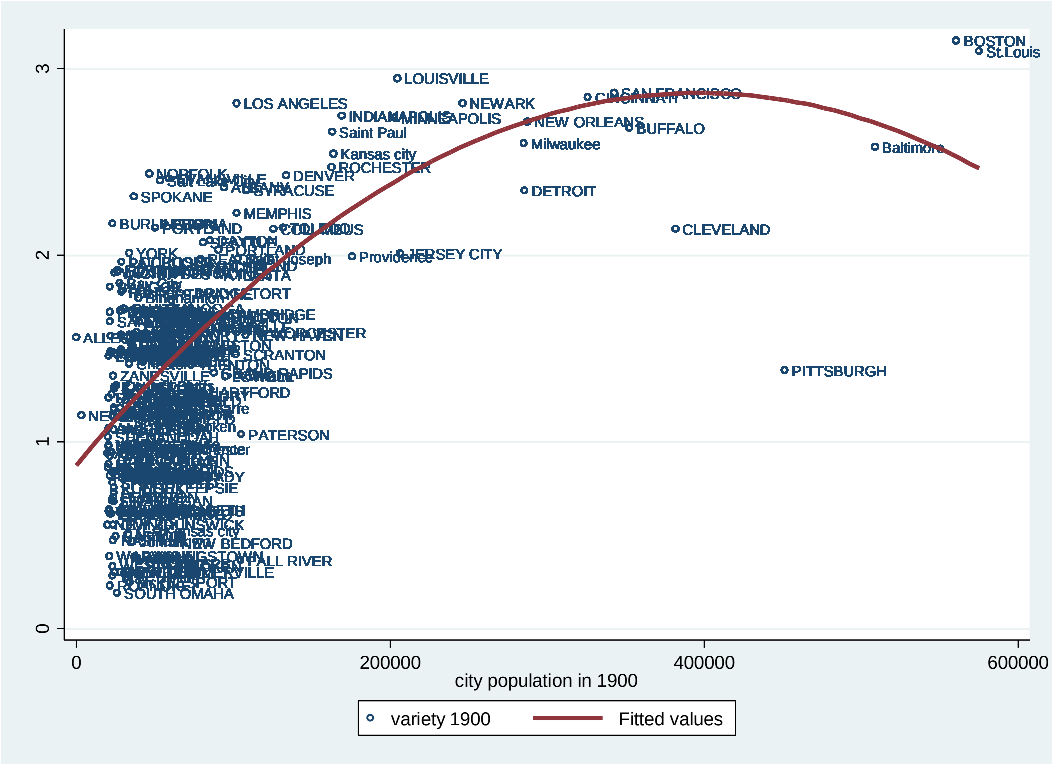
*Notes*:
1890 MAR multiplies city output by 1/1.2315; 1890 MAR and JACOBS multiplies city output by 1/1.2584.
Calculations are based on 5,423,796 production workers in cities in 1920 and assume no change in the productivity of manufacturing workers outside these cities.
*Sources*: derived from Table 11, Historical Statistics, Broadberry (1997), and Kendrick (1961)

**Figure 1: City Population vs Specialization in 1900.**




*Source*: derived from U. S. Census of Manufactures 1900.

**Figure 2: City Population vs Variety in 1900.**



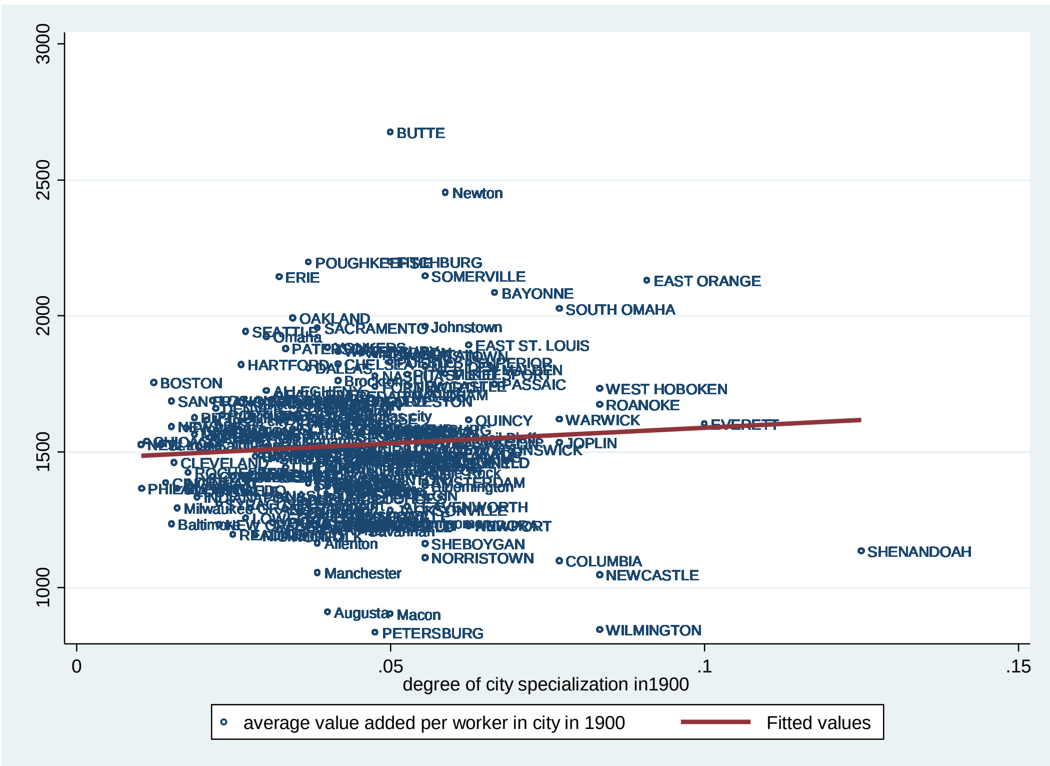

*Source*: derived from U. S. Census of Manufactures 1900.

**Figure 3. City Size, Specialization and Variety: 1890 vs. 1920.**


*Notes*: left panels: population vs. specialization; right panels: population vs. variety

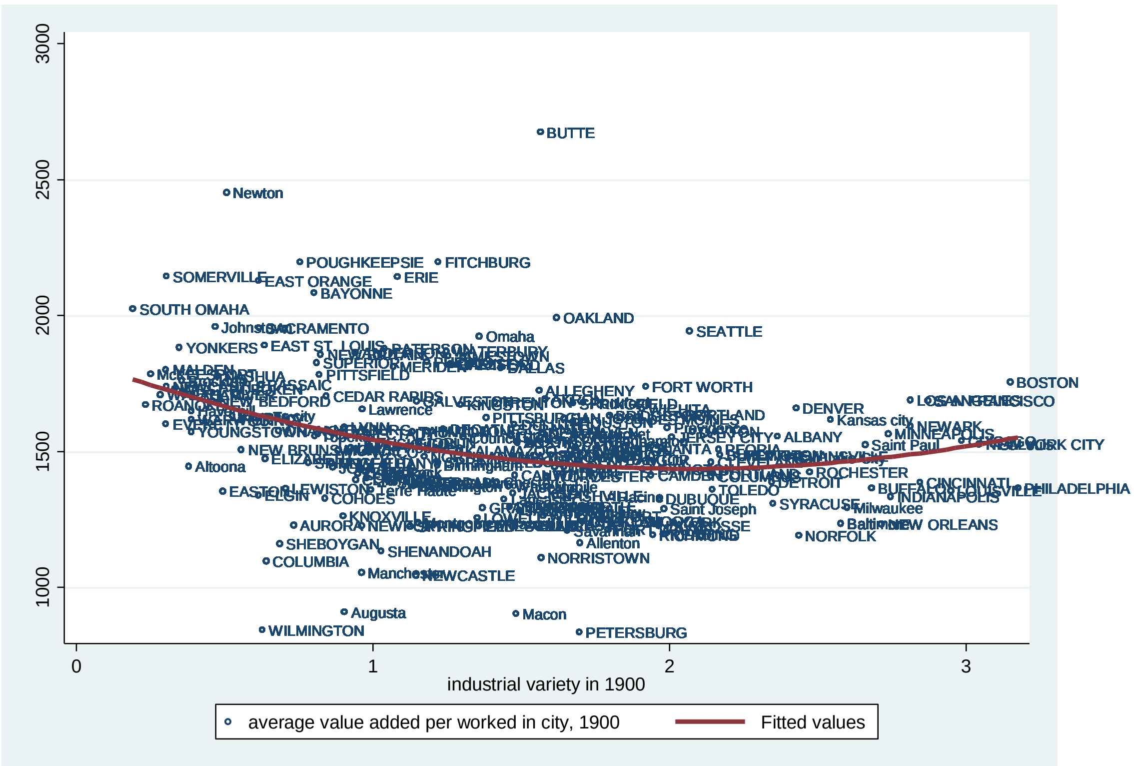
*Source*: derived from U. S. Census of Manufactures 1890 and 1920.

**Figure 4: City Labor productivity vs Specialization in 1900.**




*Source*: derived from U. S. Census of Manufactures 1900.

**Figure 5: City Labor productivity vs Variety in 1900.**




*Source*: derived from U. S. Census of Manufactures 1900.

**Figure 6: Nonparametric estimates of Jacobian externalities**