Atl Econ J (2017) 45:133–147
DOI 10.1007/s11293-017-9541-y

42 Years of Urban Growth and Industry Composition

Andrew Perumal[1]

Published online: 18 May 2017

Abstract In recent decades, knowledge spillovers have taken the helm as the driving force of growth in cities. The ease of communicating ideas and the sheer density of large urban areas have made this a plausible explanation for continued growth of employment and population in cities. However, there is little consensus on the nature of the optimal conditions for stimulating knowledge spillovers. This paper identifies these optimal conditions by exploring the relative importance of industry specialization, diversity and competition across all industries and all metropolitan areas from 1970 to 2011 in the U.S. Long-term employment growth in cities is found to be driven by industry diversity combined with a high level of competition. This combination fosters the greatest amount of cross-industry fertilization of ideas and knowledge spillovers.

Keywords Urban growth · Agglomeration · Diversity · Specialization · Competition

JEL O1 · R0 · R1

Introduction

Over the last 40 years, cities across the U.S. have had very diverse growth experiences. Nearly half of the cities that were among the top ten by population in 1970 are now not anywhere near the top of the list (Table 1), with only Houston, Los Angeles, and New York having experienced population growth during this period. This change in population levels across cities is even more dramatic if one looks even further back in time (Fig. 1). Table 1 and Fig. 1 depict an interesting milieu of an urban dynamic that has experienced numerous fluctuations in the growth and decline of population within cities over time. This naturally leads to exploring the underlying factors that govern the tumultuous decline and growth of cities.

✉ Andrew Perumal
andrew.perumal@umb.edu

[1] Department of Economics, University of Massachusetts Boston, 100 Morrissey Blvd, Boston, MA 02125, USA

Table 1 Top 10 Cities: Ranks in 1970 and 2010

Rank in 1970	City	Rank in 2010
1	New York	1
2	Los Angeles	2
3	Chicago	3
4	Philadelphia	5
5	Detroit	12
6	San Francisco	4
7	Boston	11
8	Washington, DC	10
9	Pittsburg	24
10	St. Louis	20

Source: U.S. Census Bureau (2017) and Gibson (1998).

Cities exist due to the positive externalities accruing to the spatial concentration of economic activity, that is, agglomeration economies. The exact nature of agglomeration economies, however, has been attributed to multiple sources, which have been highly contested (Andersson et al. 2014; Audretsch and Feldman 1996; Combes and Gobillon 2015; Glaeser and Gottlieb 2009; Melo et al. 2009; Puga 2010; Rosenthal and Strange 2001). The original source of agglomeration economies was the reduction of transportation costs resulting in economies of scale in production through close proximity to suppliers (Ciccone and Hall 1996; Holmes 1999; Holmes and Stevens 2002). The importance of this has declined more recently, evidenced by the shift within-cities away from manufacturing to that of service provision (Kolko 1999). Such a change in the nature of agglomeration has been linked to the rising importance in the moving of ideas and knowledge (Glaeser and Maré 2001). In this sense, dense urban areas provide a fertile landscape for the quick

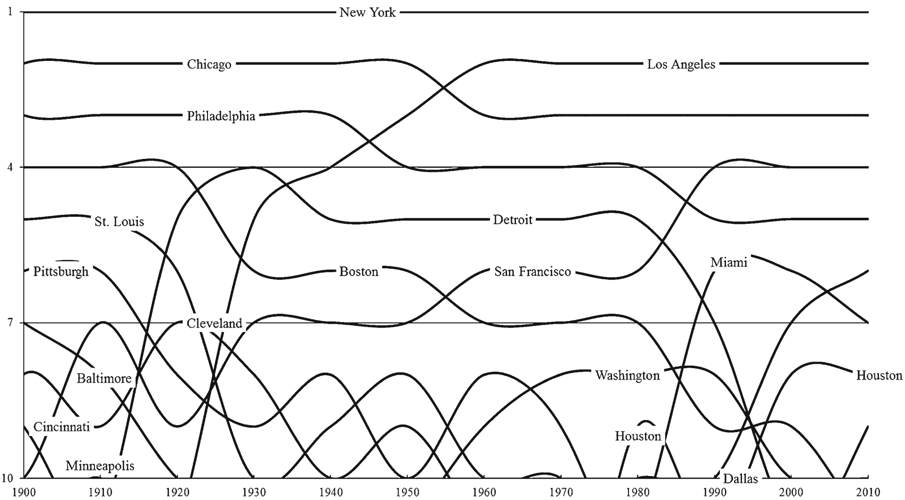

Fig. 1 Change in Rank for Top 10 Cities, 1900-2010. Sources: Peakbagger.com (2017), U.S. Census Bureau (2017), and Gibson (1998)

exchange of ideas and knowledge (Glaeser 1999; Glaeser and Maré 2001; Rauch 1993) leading to greater innovation and productivity (Moretti 2004), with the city lending itself once again to production efficiencies. Understanding agglomeration economies, therefore, sheds light on economic growth at the microeconomic level with implications for national growth. It also contributes to understanding the workings of urban labor markets (Wheeler 2001), especially given that 80% of the U.S. population resides in and nearly 85% of all jobs are located in metropolitan areas (Glaeser and Gottlieb 2009).

The means by which productivity accrues from agglomeration was first explored by Marshall (1890). Marshall separated the possible sources of agglomeration into three broad categories: input sharing, knowledge spillovers, and labor market pooling. Due to the decline of manufacturing in cities, there is greater importance placed on knowledge spillovers (Glaeser et al. 1992; Glaeser and Maré 2001), and on labor market pooling (Overman and Puga 2010; Wheeler 2001; Wheeler 2006) as the main drivers of modern agglomeration economies. While there has been substantial research into identifying the sources of agglomeration,[1] there has been much less work exploring the sources of agglomeration over time (Rosenthal and Strange 2004). Even within the research that does explore this issue (Glaeser et al. 1992; Glaeser and Maré 2001; Henderson et al. 1995; Henderson 1997) only two explore the sources of agglomeration in a dynamic setting.[2] Henderson (1997) finds lagged own-industry employment (industry specialization or concentration) to be important. Glaeser and Maré (2001) find evidence of knowledge spillovers playing a significant role.

With regard to the mechanisms by which knowledge spillovers occur, there are two main competing theories: one by Marshall-Arrow-Romer (MAR) (Marshall 1890, Arrow 1962, and Romer 1986, respectively), and the other from Jacobs (1969), with supporting theoretical work coming from Porter (1990). MAR externalities occur through knowledge spillovers within a highly concentrated industry, leading to growth of that industry. In highly concentrated industries the lack of competition allows innovative firms to accrue more of the benefits from the innovation and knowledge spillovers (Glaeser et al. 1992). Under MAR externalities, the expected channel of knowledge spillovers is a highly concentrated local industry with little or no local competition. Porter (1990) emphasized local competition, rather than the lack of it, as more supportive of industry growth by fostering more innovation and the faster adoption of innovation once it is available. Therefore, a highly-concentrated industry may be characterized by low or high competition. In contrast to MAR and Porter, Jacobs (1969) focused on industry diversity at the local level to generate greater knowledge transfers to originate from outside the industry. That is, diversity of the industry fosters cross-fertilization of innovative ideas across industries. Porter's presentation of competition is equally applicable to competition within industries which jointly create a diverse industry mix at the city-level.

This paper looks to deterministically explore the mechanisms of knowledge spillovers within a dynamic panel setting using County Business Patterns (CBP) data (U.S. Dept. of Commerce 2009, 2013) from 1970 to 2011, covering all industries and all metropolitan statistical areas (MSAs), which, to the best of my

[1] For recent reviews see Combes and Gobillon (2015), Henderson (2007), and Melo et al. (2009).
[2] Glaeser et al. (1992) and Henderson et al. (1995) explore the sources of agglomeration economies over time but not in a dynamic setting, since they only use the growth rates of employment between two points in time.

knowledge, is the first attempt to explore these issues in a comprehensive data set spanning 42 years for the U.S. To do so, three standard measures of industry composition (Combes and Gobillon 2015) are used: specialization, diversity and competition. *Specialization* is a measure of the concentration of a particular industry in a city, which accurately captures the MAR concept of externalities (Ciccone and Hall 1996; Glaeser et al. 1992; Henderson et al. 1995). That is, knowledge spillovers among closely related firms occur at a faster rate in areas characterized by concentration in that industry. *Diversity* is a measure of the potential for economic and social interaction among diverse economic sectors, which speeds up technological progress (Jacobs 1969). *Diversity* is distinct from the concept of competition since its focus is the variety of industry irrespective of the number of firms. *Competition* is a measure of the extent to which firms cannot appropriate the profits of innovation to themselves, which lowers technological progress according to MAR and increases it according to Porter (1990). These three measures capture the competing and synergistic theories that explore the dynamic nature of agglomeration economies.

This paper finds strong support for industry diversity as the driver of knowledge spillovers, which is taking place across industries, rather than within. Further, higher levels of competition aid employment growth, while industry specialization stifles it. These findings shed light on the long-term determinants of employment growth at the city-level in the U.S.

Framework

This section, based on the theoretical contributions of Glaeser et al. (1992) and Henderson et al. (1995), formalizes the setting for the empirical work that follows. Suppose that a firm in an industry, in location i at time t, chooses output, inputs, and location with the objective of maximizing profit. Its production function can be represented as:

$$Q_{it} = A_{it}(\cdot) f(N_{it}, K_{it}), \tag{1}$$

where $A_{it}(\cdot)$ is the overall level of Hicks-neutral productivity industry-wide in location i at time t, N_{it} is employment in the industry in location i at time t, and K_{it} is capital in the industry in location i at time t. Labor and capital are assumed to be mobile which allows cities to differ according to their respective level of total factor productivity. Total revenue is given by $P_{it}Q_{it}$,where P_{it} is the price of output, given by the inverse demand function $P_{it}(\cdot) = (N_{it}, MC_{it})$. $P_{it}(\cdot)$ is assumed to be decreasing in local industry output (or scale) represented by N_{it}, and where MC_{it} includes regional characteristics such as the local population and urban market centers. Total cost is given by $W_{it}N_{it} + R_{it}K_{it}$,where W_{it} is the nominal wage rate for the industry in location i at time t, and R_{it} is the price of capital for the industry in location i at time t. Firms take technology, prices, and wages, as given and maximize the objective function:

$$\text{Profit} = P_{it}Q_{it} - W_{it}N_{it} - R_{it}K_{it} \tag{2}$$

Profit maximization yields:

$$P_{it}A_{it}f_1 - W_{it} = 0, \tag{3}$$

$$P_{it}A_{it}f_2 - R_{it} = 0, \tag{4}$$

where f_1 and f_2 are the marginal products of labor and capital. Equations (3) and (4) are the standard first order conditions. Analyses of productivity hinge on whether it is Hicks-neutral or that productivity augments only some components of the production function. In line with Moomaw (1983) and Sveikauskas (1975), it is assumed that the productivity effects are Hicks-neutral.[3] This allows for focusing on eq. (3) as a means of deciphering the causes of productivity. Repositioning for dynamic externalities, eq. (3) yields the equilibrium employment level such that the local wage rate equals the value of marginal product:

$$W_{it} = A_{it}(\cdot)f_1(N_{it}; \ldots)P_{it}(\cdot). \tag{5}$$

With regard to characterizing, productivity, $A_{it}(\cdot)$, externalities are typically associated with current local own-industry scale, captured by employment, N_{it}. That is, greater industry scale fosters more opportunities for knowledge spill-overs through the concentration of people who work in the same industry interacting with each other. Furthermore, the level of technology for any industry will have both national and local components. The national level of technology is captured by national industry employment N_{it}^N differentiated by the superscript N. Therefore, firm productivity is a function of both local, A_{it}^L, and national, A_{it}^N productivity:

$$A_{it} = A_{it}^N + A_{it}^L\ . \tag{6}$$

Dynamic externalities seek to understand productivity by looking at the historical industry environment as the driving factor of current innovation and growth. This historical industry environment may be captured in many ways: own industry employment in some base period, N_{i0}^L, concentration of that employment (that is, industry *specialization*), ε_{i0}, industrial *diversity*, τ_{i0}, and *competition* in that industry, α_{i0}:

$$A_{it}^L = g\left(N_{it}^L, N_{i0}^L, \varepsilon_{it}, \tau_{it}, \alpha_{it}\right). \tag{7}$$

[3] Within the context of city growth, Hicks-neutral technology change has been a consistent approach to modeling, and measuring, productivity. Notably, Henderson (2003), using plant-level panel data for machinery and high-tech industries, identified the validity of this assumption for high-tech industries, and found only the weakest support for biased technology change in the machinery industry. Further, such explorations of Hicks-neutral technology change require plant-level firm data that includes data on capital outlays, which is not available for the breadth of industries covered in this study. Nonetheless, as the estimation approach used here focuses on employment growth, this removes the effect of time-invariant city-industry factors, including local natural advantages and the nature of industry- and city-specific technology change.

Industry-wide productivity, A_{it}, can now be characterized in terms of dynamic externalities. The growth rate of technical progress is the sum of the growth of national and local technology. Converting eq. (6) into growth rates yields:

$$log\left(\frac{A_{it+1}}{A_{it}}\right) = log\left(\frac{A^N_{it+1}}{A^N_{it}}\right) + log\left(\frac{A^L_{it+1}}{A^L_{it}}\right). \tag{8}$$

Allowing national technology to grow at a rate exogenous to local industries, yields:

$$log\left(\frac{A^N_{it+1}}{A^N_{it}}\right) = log\left(\frac{N^N_{it+1}}{N^N_{it}}\right) = \tilde{N}^N_{it}, \tag{9}$$

where the local technology is assumed to grow at a rate exogenous to the firm but dependent on the various technological externalities present in this industry in that location, using eq. (7):

$$log\left(\frac{A^L_{it+1}}{A^L_{it}}\right) = g\left(N^L_{it}, N^L_{i0}, \varepsilon_{it}, \tau_{it}, \alpha_{it}\right) + e_{it+1}. \tag{10}$$

Combining eqs. (8), (9) and (10):

$$log\left(\frac{A_{it+1}}{A_{it}}\right) = \tilde{N}^N_{it} + g\left(N^L_{it}, N^L_{i0}, \varepsilon_{it}, \tau_{it}, \alpha_{it}\right) + e_{it+1}. \tag{11}$$

Returning to eq. (5) and converting it to growth rates, yields:

$$log\left(\frac{W_{it+1}}{W_{it}}\right) = log\left(\frac{A_{it+1}}{A_{it}}\right) + log\left(\frac{f_1(N_{it+1})}{f_1(N_{it})}\right) + log\left(\frac{P_{it+1}}{P_{it}}\right), \tag{12}$$

where, using $P_{it}(\cdot) = (N_{it}, MC_{it}), log\left(\frac{P_{it+1}}{P_{it}}\right) = h(N_{it}, MC_{it})$. Using a Cobb-Douglas production function, which directly allows for Hicks-neutral productivity, of the form $Q_{it} = A_{it}K^{\beta}_{it}N^{(1-\beta)}_{it}$, where $0 < \beta < 1$, implies that $log\left(\frac{f_1(N_{it+1})}{f_1(N_{it})}\right) = \beta\, log\left(\frac{N_{it+1}}{N_{it}}\right) = \beta\, \tilde{N}^L_{it}$. Also, simplifying $log\left(\frac{W_{it+1}}{W_{it}}\right) = \tilde{W}_{it}$. Using these simplifications and eqs. (11), (12) can be rearranged for the growth of local industry employment:

$$\beta\tilde{N}^L_{it} = \tilde{W}_{it} + \tilde{N}^N_{it} + v\left(N^L_{it}, N^L_{i0}, \varepsilon_{it}, \tau_{it}, \alpha_{it}, MC_{it}\right) + e_{it+1}. \tag{13}$$

Equation (13) allows for the association of own-industry employment growth in a city with measures of technological externalities, namely specialization, competition, and diversity. In regard to defining these three measures of externalities (Combes and Gobillon 2015; Glaeser et al. 1992; Krugman 1991), higher values represent a higher degree of each attribute.

Specialization:

$$\varepsilon_i = \frac{industry\ employment\ in\ city/total\ employment\ in\ city}{industry\ employment\ in\ U.S./total\ employment\ in\ U.S.} \tag{14}$$

Diversity:

$$\tau_i = \frac{(total\ employment\ in\ city - industry\ employment\ in\ city)}{total\ employment\ in\ city} \tag{15}$$

Competition:

$$\alpha_i = \frac{firms\ in\ city{-}industry/workers\ in\ city{-}industry}{firms\ in\ U.S.\ industry/workers\ in\ U.S.\ industry} \tag{16}$$

Specialization, eq. (14), measures the extent to which a particular industry in a particular city is more or less concentrated than that industry at the national level. *Diversity*, eq. (15), measures the proportion of employment in that city that is not accounted for by a particular industry. *Competition*, eq. (16), calculates the number of firms per worker in an industry in a city and compares it to the national average of firms per worker in that industry.

The limitation of this specification is that the externalities arising from knowledge spillovers are assumed to be constant over time. Further, these externalities affect both mature and young industries equally, irrespective of product life cycle theories, whereby, younger, more innovative industries may experience the fastest growth from externalities (Glaeser et al. 1992; Henderson et al. 1995). The industry data used here does not allow for a determination of mature or young industries. The analysis also does not account for the relative spatial isolation of a particular city, which has been shown to be an important factor, possibly due to market-scale effects, in explaining population growth in cities (Dobkins and Ioannides 2001). In this vein, the estimation approach does not take into account the attenuation of spillover effects within a city (Rosenthal and Strange 2003).

Data

This paper makes use of CBP data from the U.S. Census Bureau for the years 1970–2011 (U.S. Dept. of Commerce 2009, 2013): data for 1990–2011 were obtained directly from the Census Bureau, while data for 1970–1989 was obtained from Inter-university Consortium for Political and Social Research (ICPSR) (U.S. Dept. of Commerce 2009).[4] Data were collected at the county level and aggregated to the MSA level. The CBP data

[4] The CBP data from the ICPSR are available under the following series identifiers: 24722, 8464, 8441, 8442, 8142, 8348, 8360, 8433, 8665, 8883, 9198, 9381, 9711, and 9740.

were supplemented by population data from the Census Bureau Population Estimates, and MSA-wide average years of education was obtained from the Current Population Survey (CPS) (Flood et al. 2015). At the county level, CBP data on employment is suppressed when such information would allow the identification of an individual employer, and is replaced by ranges of number of employees, and more recently through noise infusion (U.S. Census Bureau 2016). To address this, missing values for county-industry employment were replaced with the mid-point of each respective range. For example, if the reported range for employment in a particular county-industry was 100 to 249 employees, 175 was used. Even though the CBP data use ranges of employment by county-industry, when data are suppressed there is no equivalent method used for payroll information. Such data are reported as missing where employment data is suppressed, and thus payroll cannot be used in this analysis.

CBP data are currently tabulated according to the North American Industry Classification System (NAICS), while years prior to 1998 were tabulated according to the Standard Industrial Classification (SIC) System. Specifically: 2008–2011 CBP data use NAICS 2007; 2003–2007 data use NAICS 2002; 1998–2002 data use NAICS 1997; 1988–1997 data use 1987 SIC; 1973–1987 data use 1972 SIC; and 1970–1972 data use 1968 SIC. For comparability of the data over time, the respective SIC to SIC concordances for years prior to 1987 were utilized, SIC to NAICS concordances for the transition in 1997, and then NAICS to NAICS concordances thereafter. This process standardized the industry classification across the entire dataset by up-converting all industry classifications to NAICS 2007. Furthermore, the three-digit SIC and four-digit NAICS are used as the level of conversion as those levels of disaggregation refer to industry-group. This digit-level allows for a substantial amount of industry disaggregation, compared to using two-digit SIC and three-digit NAICS, as is common in the literature.

Across the sample, 29.2% of the MSA's are in the Midwest, 19.5% are in the Northeast, 43.7% are in South, and the remaining 7.6% are in the West. Average years of education increased from 10.9 to 12.7 from 1970 to 2011. Table 2 provides further descriptive statistics of the data. Between 1970 and 2011 MSA-industry employment growth averaged 1.4% per year, while national-industry employment growth averaged 2.2%. There is substantial variation in industry employment growth, bounded by −7.5% and 7.9%, and −6.1% and 6.6%, for MSA-industry and national-industry employment growth, respectively. This reflects the births, deaths, and possibly relocations, of many firms and industries since 1970. During this period, there has been substantial change in industry specialization, competition and diversity (Table 2). Industry specialization has decreased dramatically, with a substantial drop in the upper extreme of specialization. Competition has become more prevalent, but so has its extremes across the data. Finally, mean diversity has barely changed but the variation of diversity across the U.S. has doubled. Table 3 presents a correlation of key variables with region. The Midwest tends to have more specialization, less competition and industry diversity, smaller MSA populations, somewhat higher average education levels, but lower growth in MSA-industry employment. Northeastern cities are less specialized with less competition, but greater diversity, a tendency for larger cities, but a less strong correlation with average education levels, and negative MSA-industry employment growth. The South is more likely to be specialized with positive MSA-industry employment growth, less competition, diversity, population and average education. Finally, the West is more likely to be competitive, the only region that

Table 2 Descriptive statistics

	Mean	Std. Dev.	Min	Max
Log MSA Own-Industry Employment Growth	0.023	0.515	−7.484	7.847
Log National Own-Industry Employment Growth	0.035	0.292	−6.061	6.561
1970				
Specialization	5.365	19.290	0.024	1229.935
Competition	1.269	0.840	0.025	13.306
Diversity	0.997	0.005	0.868	0.999
2011				
Specialization	1.264	3.442	0.000	300.190
Competition	2.092	3.759	0.012	337.996
Diversity	0.996	0.008	0.754	0.999

Source: U.S. Dept. of Commerce, Bureau of the Census, County Business Patterns (2009, 2013), and Current Population Survey (Flood et al. 2015); Author's calculations.

exhibits this characteristic, while having less specialization, diversity, population, a higher level of average education, and positive MSA-industry employment growth.

Methodology

In the reduced form eq. (13), current local own-industry growth represents the MSA's profit opportunities to those industries, as determined by current and historical market and industry conditions (local and national), which are being acquired by firms in that MSA. *Ceteris paribus*, an improvement in the local industry environment improves profit opportunities and therefore increases local own industry scale, measured by growth in local own-industry employment. This is the process of interest, examined by estimating the following:

$$\tilde{N}_{it}^{L} = \alpha_0 + \alpha_1 \tilde{N}_{it-1}^{L} + \sum_{j=1}^{m} \delta_j X_{i,t-j} + \beta r_i + f_i + d_t + e_{it}, \tag{17}$$

Table 3 Correlations of key variables and region, across all years (1970–2011)

	Midwest	Northeast	South	West
Specialization	0.0045	−0.0147	0.0126	−0.0075
Competition	−0.0259	−0.0034	−0.0084	0.0273
Diversity	−0.0062	0.0233	−0.0124	-0.0071
MSA population	−0.0665	0.1704	−0.0748	-0.0147
MSA Average Years of Education	0.1532	0.0474	−0.1051	0.0745
Log MSA Own-Industry Employment Growth	−0.0037	−0.0075	0.0044	0.0069

All correlations are significant at the 5% level.

Source: U.S. Dept. of Commerce, Bureau of the Census, County Business Patterns (2009, 2013), and Current Population Survey (Flood et al. 2015); Author's calculations.

where $\tilde{N}_{it}^{L}$ is (log) own-industry employment growth in location i at time t. $\tilde{N}_{it-1}^{L}$ is the one-year lagged values of the dependent variable. The $X_{i,t-j}$ are time varying variables, where their lag structure starts at $t-1$ and is set to run m periods.[5] Given that growth of employment is the dependent variable, it is important to assess whether the previous years' values (at $t-1$, and deeper lags) of the independent variables affect employment growth. The independent variables, $X_{i,t-j}$, are specialization, diversity and competition, (log) national-industry employment growth,[6] as well as the (log) MSA population,[7] and average years of education.[8] The r_i are the remaining time invariant variables, which are the dummy variables for the respective regions: Midwest (omitted), Northeast, South, and West.[9] The error term is decomposed into a fixed effect,[10] f_i, a time effect, d_t, applying to all localities in time t, and a contemporaneous drawing e_{it}. f_i represents the influence of time invariant unobserved and unmeasured characteristics of the local area, controlling for this prevents location-specific omitted variables from biasing the estimation. The e_{it} are generally assumed to be i.i.d. across space and time.

Estimating eq. (17) presents the problem that the fixed effects, f_i, are correlated with other right-hand side variables, thereby biasing ordinary least squares estimates. Estimation of eq. (17) with standard fixed effects procedures still results in biased estimates because the contemporaneous error term, e_{it}, is correlated with any time average of the lagged dependent variable, $\tilde{N}_{it-1}^{L}$. Accordingly, it is necessary to take first differences of eq. (17) to obtain:

$$\Delta\tilde{N}_{it}^{L} = \alpha_1\Delta\tilde{N}_{it-1}^{L} + \sum_{j=0,1}^{m} \delta_j \Delta X_{i,t-j} + \Delta d_t + \Delta e_{it}, \tag{18}$$

where $\Delta\tilde{N}_{it}^{L} = \tilde{N}_{it-j}^{L} - \tilde{N}_{it-j-1}^{L}$ and $\Delta e_{i,t} = e_{i,t-j} - e_{i,t-j-1}$. As seen in eq. (18), first differencing removes all the time invariant variables, namely the regional and fixed/location effects dummies. First differencing of eq. (17) to obtain eq. (18), while solving the endogeneity problem, does introduce by construction serial correlation since $\Delta\tilde{N}_{it-1}^{L}$ and Δe_{it-1} are correlated with Δe_{it}. To correctly account for this, estimation is conducted by two-step feasible generalized method of moments (GMM). For instrumentation,[11] predetermined values of the lagged dependent variable are used, specifically lags of order three through five.[12] The use of two-step feasible GMM provides efficiency gains, relative to two-stage least squares, in the presence of over identifying restrictions (Hayashi 2000, pp. 204–235). Variance estimates are heteroskedastic and autocorrelation-consistent (HAC) through the use of Newey-West kernel estimation, with a kernel bandwidth of three, selected by T^(1/3) = 42^(1/3) = 3.476. The GMM weighting matrix has been centered so as to have a mean of zero as the finite-sample performance is improved (Hall

[5] Computational limitations restrict the lagged values of the explanatory variables to five periods.

[6] Controls for industry-wide technology innovation and fluctuations in market demand for industry output.

[7] Controls for the effects of the size of the urban area on employment growth.

[8] Controls for the impact of having a highly educated population living and working in metropolitan areas.

[9] Captures unobserved year- and region-specific idiosyncrasies that may contribute to growth in that year or region.

[10] Based on the Hausman specification test.

[11] Combes et al. (2011), Combes et al. (2012), and Combes and Gobillon (2015), provide an in-depth review.

[12] Use of the 1st and 2nd lags of the dependent variable resulted in not rejecting the null of the Hansen J test.

2005, pp. 131–8, 145–8). The estimation results pass tests of under- and weak-identification.[13] All the results presented here do not reject the null hypothesis of valid instruments (Hansen J test of overidentifying restrictions). That is, the instruments are uncorrelated with the error term, and the excluded instruments are correctly excluded from the estimated equation. Testing the endogeneity of the endogenous regressor, using the GMM distance statistic, proves that the lagged dependent variable is not exogenous.

Results

Table 4 presents the results for the two-step feasible GMM estimation of eq. (18). The results strongly support Jacob's industry diversity and Porter's competition between firms as the determining factors of long-term employment growth. The results show that a one-year lag of own-industry employment growth positively affects current year employment growth, where a 1% growth in a particular industry in a MSA in the previous year results in a 0.18% growth in that industry this year, reflecting the importance of own-industry inter-temporal spillovers. Industry *Specialization* is found to have little to no effect on employment growth, except from the immediately preceding year, where a one-point increase in the industrial specialization in the preceding year, reduces employment growth by 6.1%. Industry concentration does not lead to innovation and growth. Industry diversity is the vital determinant of employment growth with positive and significant coefficients across four of the five-year lags. The strongest effect of diversity on employment growth occurs in the immediately preceding year: a 0.01-point increase[14] in industrial diversity results in an increase of 81.9% in employment growth.[15] Competition contributes positively to employment growth with significant coefficients throughout the five-year lags: in the immediately preceding year, a one-point increase in competition results in a 9.1% increase in employment growth. These findings indicate that a diversified industry base bolstered with competition fosters growth by means of cross-industry knowledge spillovers and quick adoption of innovations.

National own-industry employment growth controls for industry-wide technology innovation and fluctuations in market demand, both of which would affect own-industry employment growth at the local level. A 1% increase in national own-industry growth results in a 0.76% decline in industry growth at the MSA level. MSA population controls for the effects of the size of the urban area on employment growth, and is significant and negative at lag 1, small and positive in lag 2, and insignificant thereafter. Therefore, increases in MSA population do not beneficially affect employment growth in any particular local industry, possibly due to the congestion costs of doing business in locations with an increasing urban population. Finally, average years of education have no impact on employment growth across all the five lags. This finding contrasts with the literature on this

[13] Use of HAC variance estimates necessitate Kleibergen-Paap rk, and Kleibergen-Paap Wald rk F statistics.

[14] The measure of industry diversity is bounded between 0 and 1. Therefore, a typical one-point increase interpretation of the coefficient is incorrect for this variable.

[15] The magnitude of this effect is noticeably large. However, alternate specifications using only a one-year lag of industrial diversity yields a coefficient of 79.91. Results are available upon request.

Table 4 Two-step feasible GMM estimation results, 1970–2011

	Coef.	Std. Err.
Log Own-Industry Employment Growth, 1-yr. lag	0. 18***	(0.040)
Specialization 1-yr. lag	−0. 061***	(0.0088)
2-yr. lag	0. 0053	(0.0038)
3-yr. lag	0. 0004	(0.0019)
4-yr. lag	0. 0007	(0.0017)
5-yr. lag	0.0015	(0.001)
Diversity 1-yr. lag	81.9***	(6.81)
2-yr. lag	−5.61	(3.47)
3-yr. lag	6.18***	(1.12)
4-yr. lag	8.50***	(1.08)
5-yr. lag	9.01***	(1.05)
Competition 1-yr. lag	0. 091***	(0.012)
2-yr. lag	0. 0087	(0.0047)
3-yr. lag	0.013***	(0.0023)
4-yr. lag	0.0059***	(0.0015)
5-yr. lag	0.0038**	(0.0012)
Log National Own-Industry Employment Growth 1-yr. lag	−0.76***	(0.032)
2-yr. lag	−0.50***	(0.010)
3-yr. lag	−0.36***	(0.007)
4-yr. lag	−0.24***	(0.0064)
5-yr. lag	−0.11***	(0.0047)
Log MSA Population 1-yr. lag	−0.46***	(0.038)
2-yr. lag	0.049*	(0.023)
3-yr. lag	0.0058	(0.021)
4-yr. lag	0.014	(0.021)
5-yr. lag	0.0051	(0.018)
MSA Average Years of Education 1-yr. lag	0.0029	(0.0035)
2-yr. lag	−0.0033	(0.0036)
3-yr. lag	−0.0017	(0.0037)
4-yr. lag	−0.0040	(0.0033)
5-yr. lag	0.00034	(0.0034)
Constant	0.020***	(0.0033)
Observations: 474,098; Adjusted R^2: 0.058		

***, **, * denote significance at the 0.1%, 1% and 5% level, respectively. Controls include dummy variables for years and regions. Identification test statistics: Weak identification test (Kleibergen-Paap rk Wald F): 103.352; Underidentification test (Kleibergen-Paap rk LM): 212.167, p = 0.0000; Over-identification test (Hansen J): 2.161, p = 0.3394. Source: U.S. Dept. of Commerce, Bureau of the Census, County Business Patterns (2009, 2013), and Current Population Survey (Flood et al. 2015); Author's calculations.

issue, which finds a positive relationship between education and employment growth. However, that literature typically explores aggregate employment growth rather than unweighted growth across individual industries.

The results regarding industry specialization versus diversity are in stark contrast to the majority of research in the literature that explored this issue for shorter time periods, and for a limited number of industries. Many studies have repeatedly found that specialization matters more (Henderson 2003; Henderson et al. 1995; Henderson 1997; Holmes and Stevens 2002). However, the findings presented here supersede those findings in two important ways. First, this paper provides a comprehensive analysis of industries at the four-digit NAICS (and three-digit SIC) level over an extended period of time. To the best of my knowledge, no other research has performed this type of analysis on such a comprehensive data set for the U.S. Further, the majority of the literature focuses on static externalities. This paper joins a small collection of research endeavors that explore these issues in a dynamic setting.

Conclusion

Industry diversity is clearly important in driving employment growth, stemming from knowledge spillovers that are occurring across industries. Competition also drives employment growth, possibly by encouraging the rapid adoption of new technology. In combination, these factors hold the most value in explaining employment growth over a long period of time.

Understanding the dynamic mechanisms of knowledge spillovers guides policy decisions regarding industry composition. Previous research into dynamic externalities has favored specialization over diversity as the key determinant in employment growth. However, that research is limited to relatively short periods of time, infrequent time points, and limited industry coverage. This paper overcomes these shortfalls by covering all industries and all MSAs over the 42-year period beginning in 1970. These results, therefore, hold the most potential for broad applicability in understanding dynamic industry composition and employment growth across cities in the U.S. These findings also mirror those of similar studies which explored this issue for England (Hanlon and Miscio 2014) and France (Combes et al. 2012). Most importantly, the findings presented here should spur a careful exploration of the long-term benefits of local and regional development policies (Kline: and Moretti 2013), which have typically focused on industry-specific growth, rather than seeking to enhance industry diversity or competition.

References

Andersson, M., Klaesson, J., & Larsson, J. P. (2014). The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies? *Papers in Regional Science, 93*(4), 727–747.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies, 29*(3), 155–173.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the Geography of innovation and production. *The American Economic Review, 86*(3), 630–640.

Ciccone, A., & Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review, 86*(1), 54–70.

Combes, P., Gobillon, L., (2015). Chapter 5 - The Empirics of Agglomeration Economies. In Duraton, G., Henderson, J.V., Strange, W.C. (Eds.), *Handbook of Regional and Urban Economics vol. 5* (pp. 247–348). Elsevier.

Combes, P., Duranton, G., & Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography, 11*(2), 253–266.
Combes, P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica, 80*(6), 2543–2594.
Dobkins, L. H., & Ioannides, Y. (2001). Spatial interactions among U.S. Cities: 1900–1990. *Regional Science and Urban Economics, 31*(6), 701–731.
Flood, S., King, M., Ruggles, S., Warren, J.R., 2015. Integrated public use microdata series, current population Survey: Version 4.0. (machine-readable database). Minneapolis: University of Minnesota. Accessed 1 Feb 2016.
Gibson, C. (1998). Population of the 100 Largest Cities and Other Urban Places in The United States: 1790 To 1990. https://www.census.gov/population/www/documentation/twps0027/twps0027.html. Accessed 27 Apr 2017.
Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics, 46*(2), 254–277.
Glaeser, E.L., Gottlieb, J.D., (2009). The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. NBER Working Paper 14806.
Glaeser, E. L., & Maré, D. (2001). Cities and skills. *Journal of Labor Economics, 19*(2), 316–342.
Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy, 100*(6), 1126.
Hall, A. R. (2005). *Generalized method of moments*. New York: Oxford University Press.
Hanlon, W., & Miscio, A. (2014). *Agglomeration: A dynamic approach*. No. 20728. National Bureau of Economic Research.
Hayashi, F. (2000). *Econometrics*. Princeton: Princeton University Press.
Henderson, J. V. (1997). Externalities and industrial development. *Journal of Urban Economics, 42*(3), 449–470.
Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics, 53*(1), 1–28.
Henderson, J. V. (2007). Understanding knowledge spillovers. *Regional Science and Urban Economics, 37*(4), 497–508.
Henderson, J. V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy, 103*(5), 1067.
Holmes, T. J. (1999). Localization of industry and vertical disintegration. *The Review of Economics and Statistics, 81*(2), 314–325.
Holmes, T. J., & Stevens, J. J. (2002). Geographic concentration and establishment scale. *The Review of Economics and Statistics, 84*(4), 682–690.
Jacobs, J., 1969. The Economy of cities. Vintage, New York.
Kline, P. M., & Moretti, E. (2013). *Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley authority.* No. 19293. National Bureau of Economic Research.
Kolko, J. (1999). Can I Get Some Service Here? Information Technology, Service Industries, and the Future of Cities, SSRN eLibrary.
Krugman, P. (1991). Increasing returns and economic Geography. *Journal of Political Economy, 99*(3), 483–499.
Marshall, A. (1890). *Principles of Economics*. London: Macmillan.
Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics, 39*(3), 332–342.
Moomaw, R. L. (1983). Is population scale a worthless surrogate for business agglomeration economies? *Regional Science and Urban Economics, 13*(4), 525–545.
Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review, 94*(3), 656–690.
Overman, H.G., Puga, D. (2010). Labor pooling as a source of agglomeration: An empirical investigation. In *Anonymous Agglomeration Economics* (pp. 133–150). University of Chicago Press.
Peakbagger.com. (2017). Graph of Metro Area Population Rank over Time. http://www.peakbagger.com/pbgeog/histmetropop.aspx#method. Accessed 27 Apr 2017.
Porter, M.E., 1990. The competitive advantage of nations. Free Press, New York.
Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science, 50*(1), 203–219.
Rauch, J. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics, 34*(3), 380–400.
Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy, 94*(5), 1037.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics, 50*(2), 191–229.

Rosenthal, S. S., & Strange, W. C. (2003). Geography, industrial organization, and agglomeration. *The Review of Economics and Statistics, 85*(2), 377–393.

Rosenthal, S.S., Strange, W.C. (2004). Chapter 49 Evidence on the nature and sources of agglomeration economies. In Vernon Henderson, J., Thisse, J.-F., (Ed.), *Handbook of Regional and Urban Economics, vol. 4* (pp. 2119–2171). Elsevier.

Sveikauskas, L. (1975). The productivity of cities. *The Quarterly Journal of Economics, 89*(3), 393–413.

U. S. Census Bureau (2016). Data Withheld from Publication. http://www.census.gov/programs-surveys/cbp/technical-documentation/methodology/data-witheld-from-publication.html. Accessed 13 June 2016.

U.S. Census Bureau (2017). Population and Housing Unit Estimates. https://www.census.gov/programssurveys/popest/data/tables.html. Accessed May 5, 2017.

U.S. Dept. of Commerce, Bureau of the Census (2009). County Business Patterns (United States): U.S. Summary, State and County Data. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census (producer). Ann Arbor: Inter-university Consortium for Political and Social Research (distributor). https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/00022. Accessed 11 Jan 2016.

U.S. Dept. of Commerce, Bureau of the Census (2013). County Business Patterns (United States): U.S. Summary, State and County Data. https://www.census.gov/programs-surveys/cbp.html. Accessed 18 Jan 2016.

Wheeler, C. H. (2001). Search, sorting, and urban agglomeration. *Journal of Labor Economics, 19*(4), 879–899.

Wheeler, C. H. (2006). Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics, 60*(2), 162–184.