*Urban production externalities**

Antonio Ciccone,
ICREA and UPF

November 2006
Preliminary — Comments Welcome

Abstract:
Urban production externalities (agglomeration effects) are external effects among producers in areas with a high density of economic activity. Such external effects are the main explanation for why productivity is usually highest in those areas of a country where economic activity is densest. There is some disagreement about the strength of urban production externalities. What is clear however is that even weak urban production externalities can explain large spatial differences in productivity.

* Prepared for *New Palgrave: A Dictionary of Economics*

Author's address:
Department of Economics and Business
UPF
Ramon Trias Fargas 25-27
08005 Barcelona, Spain
antonio.ciccone@upf.edu
http://www.crei.cat/people/ciccone/welcome.html

1. Introduction

Urban production externalities (agglomeration effects) are external effects among producers in areas with a high density of economic activity. Such external effects are the main explanation for why productivity is usually highest in those areas of a country where economic activity is densest. The best understood urban production externalities are technological externalities due to knowledge spillovers and non-transportable input sharing, both of which are already discussed by Marshall (1920).

That local technological externalities translate into increasing returns at the city level is demonstrated formally by Henderson (1974). Building on the analysis of Chipman (1970), he also shows that such increasing returns are compatible with perfect competition. Abdel-Rahman (1988), Fujita (1988, 1989), and Rivera-Batiz (1988) present a rigorous analysis of decentralized market equilibria with increasing returns at the city level due to intermediate input sharing. These contributions build on the formalization of monopolistic competition of Spence (1976) and Dixit and Stiglitz (1977) to show how increasing returns to city size emerge when some intermediate inputs are non-transportable and produced subject to increasing returns at the plant level.

There is some disagreement about the strength of increasing returns to the density (or scale) of local economic activity and therefore about the strength of urban production externalities. This is partly because the best approach to estimation is still unclear. What is clear however is that even weak urban production externalities can explain much of the large spatial differences in productivity observed in many countries. This is because spatial differences in the density of economic activity are very large so that even a small degree of increasing returns to density can explain sizable spatial productivity differences. Moreover, mobile physical capital and tradable intermediate inputs imply that the strength of increasing returns to density exceeds the strength of urban production externalities (by approximately a factor of two).

The remainder of this entry first illustrates the link between the strength of urban production externalities and the strength of increasing returns to the density of economic activity (or increasing returns at the city level). It then turns to the advantages and drawbacks of different empirical approaches to urban production externalities.

2. A Model of Urban Production Externalities and Increasing Returns

The link between urban production externalities and increasing returns to the density of economic activity is easily illustrated using the technology-spillover model of Ciccone and Hall (1996) extended to include costlessly tradable intermediate inputs. This extension is important for understanding why the strength of increasing returns to density is approximately twice the strength of urban production externalities. Including urban production externalities due to non-transportable intermediate-input sharing in the model would be straightforward, see Ciccone and Hall, but not change any of the relevant conclusions.

Model Setup

Let $f(n_f, k_f, m_f; Q_c, A_c)$ be the production function that describes the amount of output produced by firm f on an acre of space when employing n workers, m units of costlessly tradable intermediate inputs, and k units of capital (small letters denote per-acre quantities). The acre is embedded in county c with total output Q and total acreage A (capital letters denote total county-level quantities). The simplest production function to deal with is one where the externality depends multiplicatively on the density of economic activity Q/A, and the elasticity of $f(n, k, m; Q, A)$ with respect to all its arguments is constant. In this case,

(1) $$q_f = \left(n_f^{\alpha} k_f^{\beta} m_f^{1-\alpha-\beta}\right)^{1-\rho} \left(\frac{Q_c}{A_c}\right)^{\lambda}.$$

$\lambda \geq 0$ captures the strength of urban production externalities (agglomeration effects); for example, $\lambda = 2\%$ means that a doubling of the density of economic activity is associated with a 2% increase in the output of the firm (for a given amount of inputs used by the firm). $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$ determine the relative importance of labor, capital, and intermediate inputs in production. And $0 \leq \rho \leq 1$ captures possible decreasing returns to labor, capital, and intermediate inputs when holding the amount of land used in production constant (congestion effects).

Input Demand and Value Added

Firms maximize profits taking factor prices and aggregate output in each county as given. Profit maximization implies that firms employ capital up to the point where its marginal product is equal to the national rental price of capital R (measured in units of output), which gives rise to a demand for capital equal to $k_f = \beta(1-\rho)q_f / R$. The demand for intermediate inputs can be determined analogously as $m_f = (1-\alpha-\beta)(1-\rho)q_f$, where we have assumed that one unit of intermediate inputs can be produced with one unit of output. Substituting these factor demand functions in (1) and solving for output at the firm level yields that q_f is proportional to $\left(n_f\right)^{\alpha(1-\rho)/(1-(1-\rho)(1-\alpha))}\left(Q_c / A_c\right)^{\lambda/(1-(1-\rho)(1-\alpha))}$.

Moreover, the demand for intermediate inputs implies that value added at the firm level (y) and county level (Y) are a fraction $1-(1-\alpha-\beta)(1-\rho)$ of the total value of production at the firm and county level respectively, i.e.

$y_f = q_f - m_f = \left(1-(1-\alpha-\beta)(1-\rho)\right)q_f$ and $Y_c = Q_c - M_c = \left(1-(1-\alpha-\beta)(1-\rho)\right)Q_c$.

Hence, value added at the firm level is linked to firm-level employment and county-level value added by

$$(2) \qquad y_f = \left(\gamma n_f\right)^{\frac{\alpha(1-\rho)}{1-(1-\rho)(1-\alpha)}}\left(\frac{Y_c}{A_c}\right)^{\frac{\lambda}{1-(1-\rho)(1-\alpha)}},$$

where γ is an unimportant constant.

Increasing Returns to Density

The amount of labor N employed in a county is taken to be distributed uniformly in space; $n_f = N_c / A_c$ for all firms f in the county. Substituted in (2), this yields that output per acre in a county is linked to employment per acre by

$$(3) \qquad \frac{Y_c}{A_c} = \left(\gamma \frac{N_c}{A_c}\right)^{1+\theta},$$

where the strength of increasing returns to the density of economic activity θ is given by

$$(4) \qquad \theta = \frac{(\lambda-\rho)}{\alpha(1-\rho)-(\lambda-\rho)}.$$

As expected, increasing returns to density are stronger when urban production externalities λ are strong and congestion effects ρ are weak. A necessary condition for productivity to be greater in areas with dense economic activity is that urban production externalities (agglomeration effects) more than offset congestion effects, $\theta > 0$.

From Increasing Returns to Density to Urban Production Externalities

The relationship between increasing returns to the density of economic activity θ and the strength of net agglomeration effects $\lambda - \rho$ in (4) depends on $\alpha(1-\rho)$, the elasticity of output with respect to labor. In equilibrium, this elasticity equals the share of labor in the total value of production. In the US, the share of labor in value added is around 2/3 (e.g. Gollin, 2002) and the share of intermediate inputs in value added around 1/2 (e.g. Basu, 1995), which implies a share of labor in the total value of production around 1/3. To see that this implies that urban production externalities are magnified, note that for small values of $\lambda - \rho$ (4) implies

(5) $$\theta \cong \frac{\lambda - \rho}{\alpha(1-\rho)} = 3\left(\lambda - \rho\right),$$

where we have made use of $\alpha(1-\rho) = 1/3$. A one-point increase in the strength of urban production externalities therefore implies a three-point increase in the strength of increasing returns to the density of economic activity. Much of this magnification is due to the presence of intermediate inputs. In a model without intermediate inputs where physical capital earns 1/3 of value added, the factor of magnification would have been (only) 3/2.

3. Empirical Approaches and Results

Increasing Returns to City or Industry Size

Early empirical studies of urban scale effects by Sveikauskas (1975), Segal (1976), Moomaw (1981, 1985), and Tabuchi (1986) focused on the link between city size and productivity at the city and the city-industry level. The empirical results indicate that doubling city size increases productivity by 3 to 8%. Nakamura (1985) and Henderson (1986, 2003) extend the analysis by distinguishing between urbanization economies and localization economies. Localization economies are increasing returns related to the size of city-industries, while urbanization economies refer to increasing returns to overall city size. Henderson concludes that scale effects are mostly at the industry level, but Nakamura finds evidence for both urbanization and localization economies.

Most studies of the strength of agglomeration economies measure output as the value of production or value added from the *Census of Manufacturers*. This data set does not contain information on the value of services that plants purchase in the market or obtain from headquarters. Census of Manufacturers data will therefore overstate the value added of city-industries. This bias is likely to be greater in larger cities for two reasons. First, there is more service outsourcing in larger cities, due to the larger variety of services available (Holmes, 1999; Ono, forthcoming). Second, headquarter services are more likely to be counted twice in larger cities, as such cities are more likely to contain both a plant and its headquarters. The total value of production from the Census of Manufacturers has the additional disadvantage of counting twice all intermediate inputs traded within and across industries located in the same city.

Increasing Returns to Density and the Productivity of US States

In the US, the finest level of geographical detail with reliable data on value added is the state level. Ciccone and Hall (1996) therefore estimate increasing returns to the density of economic activity by combining state-level value added data with the model in (3). Aggregating county-level value added to the state level yields that labor productivity in state s, Y_s / N_s, is equal to

(6) $$\frac{Y_s}{N_s} = D_c(\theta) \equiv \sum_{c=1}^{C_s} \left(\frac{\gamma N_c}{A_c} \right)^{1+\theta} \frac{N_c}{N_s},$$

where C_s is the number of counties in the state. Hence, the strength of increasing returns to the county-level density of economic activity can be estimated by combining cross-state variation in labor productivity and the state-level density index $D_c(\theta)$, which depends on county-level employment density and the distribution of employment across counties. Ciccone and Hall find θ to be just above 5% using a least-squares approach. Because of large differences in the density of economic activity, this limited degree of increasing returns to density can explain more than half of the sizable differences in output per worker across US states.

Ciccone and Hall's work is about the degree of increasing returns to the density of economic activity, not about the strength of urban production externalities. Going from one to the other is rather straightforward however. Using (5) yields that θ equal to 5% corresponds to a net agglomeration effect $\lambda - \rho$ of 1.7%. According to the *Flow of Funds Accounts of the United States, 1982-1990* prepared by the *Board of Governors of the Federal Reserve System* (1997), the share of land in the total value of production ρ in the private sector outside of agriculture and mining is around 0.5%. Hence, λ is between 2 and 2.5%, which implies that a doubling of the density of economic activity in a county is associated with a 2 to 2.5% increase in the output firms produce with a given amount of inputs (see (1)). Mobile physical capital and tradable intermediate inputs therefore imply that the strength of increasing returns to density exceeds the strength of urban production externalities by a factor of two. Hence, more than half of the differences in output per worker across US states can be explained by rather weak urban production externalities.

An important concern when estimating agglomeration economies is potential feedback from productivity to the density of economic activity. To address this possibility, Ciccone and Hall (1996) use an instrumental variables approach. The instruments for the state-level density index used are the population and population density of US states between 1850 and 1880, as well as the presence or absence of a railroad in each state in 1860 and the distance of states from the eastern seaboard. The identifying assumption is that the original sources of agglomeration in the US have remaining influences only through the preferences of people about where to live. The instrumental variables estimates of θ are between 5.5 and 6.1%, and therefore very similar to the least squares estimates.

Agglomeration Effects in Europe

For many European countries there is value added data at a much finer level of geographic detail than for the US. Employing such data for France, Germany, Italy, Spain, and the UK, Ciccone (2002) finds an average degree of increasing returns to the local density of economic activity between 4 and 5%, only slightly below estimates for the US. Rice, Venables, and Patacchini (2006) find a similar result using geographically detailed earnings data for the UK. They also take into account the scale of production in neighboring locations weighted by travel times and find that productivity benefits diminish quickly with travel distance.

Human Capital Externalities?

An open question is whether there are agglomeration economies associated with the geographic concentration of human capital. Rauch (1993) examines this issue by augmenting a standard Mincerian wage regression (e.g. Card, 1999) with data on the characteristics of cities where people live. His empirical model relates wages of individuals i in cities c, w_{ic}, to relevant individual characteristics (e.g. education, experience), X_{ic}, and to the average level of schooling of the city, S_c, and other city characteristics, Z_c,

$$(7) \qquad \log w_{ic} = aX_{ic} + bS_c + cZ_c + \varepsilon_{ic},$$

where ε_{ic} summarizes all other (unobserved) factors affecting individual wages across cities. Least-squares estimation of (7) using US data for 1980 yields a positive and significant coefficient on average schooling in the city (b), and Rauch therefore concludes that there are human capital externalities at the city level.

A drawback of Rauch's approach is that it cannot account for time-invariant unobserved city characteristics that increase both schooling and wages. Another drawback is that city-level schooling is taken to be exogenous. Acemoglu and Angrist (2001) address these drawbacks by taking US states to be the relevant aggregate unit in (7) (instead of cities). In this case, the data allow for an analysis of the effects of increases in average state-level schooling on individual wages. Moreover, Acemoglu and Angrist show that changes in average schooling at the state level can be instrumented by state-level changes in compulsory-schooling and child-labor laws. Their approach yields no evidence of significant schooling externalities between 1960 and 1980.

Ciccone and Peri (2006) show that a positive effect of average schooling in a Mincerian wage equation like (7) may not reflect human capital externalities but a downward sloping demand function for human capital. They therefore propose an alternative approach, which exploits that the wage differential between more and less educated workers reflects differences in marginal social products of the two worker types when human capital externalities are absent. This approach does not yield evidence for significant human capital externalities at the level of US cities or states between 1960 and 1990.

Moretti (2004) finds that US cities where the labor force share of college graduates increased most between 1980 and 1990 also saw the largest wage increase for college graduates. Using Census of Manufacturers plant-level data, Moretti (2006) finds that the output of plants in high-tech city-industries rises with levels of schooling in other high-tech industries in the same city. This evidence is consistent with human capital externalities. An alternative explanation could be that skill-biased technological progress translated into increases in the productivity and wages of college graduates in high-tech industries. Cities specialized in industries experiencing rapid productivity growth would in this case see faster output growth and attract more college graduates. This alternative hypothesis is especially plausible for the 1980-1990 period, which was marked by rising college wage premia due to skill-biased technological progress (e.g. Katz and Murphy, 1992).

References

H. M. Abdel-Rahman. Product differentiation, monopolistic competition and city size. *Regional Science and Urban Economics*, 18(1):69–86, 1988.

D. Acemoglu and J. Angrist. How large are the social returns to education: Evidence from compulsory schooling laws. In B. Bernanke and K. Rogoff, editors, *NBER Macroeconomic Annual 2000*, pages 9–59. The MIT Press, 2001.

S. Basu. Intermediate goods and business cycles: Implications for productivity and welfare. *American Economic Review*, 85(3):512-531, 1995.

Board of Governors of the Federal Reserve System. *Flow of Funds Accounts of the United States, 1982-1990*. Washington DC, 1997.

D. Card. The causal effect of education on earnings. In O. Ashenfelter and D. Card, editors, *Handbook of labor economics*. Amsterdam: North Holland, 1999.

J. S. Chipman. External economies of scale and competitive equilibrium. *Quarterly Journal of Economics*, 84(3):347–385, 1970.

A. Ciccone. Agglomeration effects in Europe. *European Economic Review*, 46(2):213–227, 2002.

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

A. Ciccone and G. Peri. Identifying Human Capital Externalities: Theory with Applications. *The Review of Economic Studies*, 73(2):381–412, 2006.

A. K. Dixit and J. E. Stiglitz. Monopolistic competition and optimum product diversity. *The American Economic Review*, 67(3):297–308, 1977.

M. Fujita. A monopolistic competition model of spatial agglomeration: Differentiated product approach. *Regional Science and Urban Economics*, 18(1):87–124, 1988.

M. Fujita. *Urban economic theory: Land use and city size*. Cambridge: Cambridge University Press, 1989.

D. Gollin. Getting income shares right. *Journal of Political Economy*, 110(2):458-474, 2002.

J. V. Henderson. The sizes and types of cities. *The American Economic Review*, 64(4):640–656, 1974.

J. V. Henderson. Efficiency of resource usage and city size. *Journal of Urban Economics*, 19(1):47–70, 1986.

J. V. Henderson. Marshall's scale economies. *Journal of Urban Economics*, 53(1):1–28, 2003.

T. Holmes. Localization of industry and vertical disintegration. *Review of Economics and Statistics*, 81(3):314-325, 1999.

L. F. Katz, and K.M. Murphy. Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–78, 1992.

A. Marshall. *Principles of economics*. London: MacMillan, 8th edition, 1920.

R. L. Moomaw. Productivity and city size: A critique of the evidence. *The Quarterly Journal of Economics*, 96(4):675–688, 1981.

R. L. Moomaw. Firm location and city size: Reduced productivity advantages as a factor in the decline of manufacturing in urban areas. *Journal of Urban Economics*, 17(1):73–89, 1985.

E. Moretti. Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1-2):175–212, 2004.

E. Moretti. Workers education, spillovers, and productivity: Evidence from plant-level production functions. *The American Economic Review*, 94:656–690(35), 2004.

R. Nakamura. Agglomeration economies in urban manufacturing industries: A case of Japanese cities. *Journal of Urban Economics*, 17(1):108–124, 1985.

Y. Ono. Market thickness and outsourcing services. *Regional Science and Urban Economics*, forthcoming.

P. Rice, A. J. Venables, and E. Patacchini. Spatial determinants of productivity: Analysis for the regions of Great Britain. *Regional Science and Urban Economics*, 36(6):727–

752, 2006.

J. E. Rauch. Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics*, 34(3):380–400, 1993.

F. L. Rivera-Batiz. Increasing returns, monopolistic competition, and agglomeration economies in consumption and production. *Regional Science and Urban Economics*, 18(1):125–153, 1988.

D. Segal. Are there returns to scale in city size? *The Review of Economics and Statistics*, 58(3):339–350, 1976.

M. Spence. Product selection, fixed costs, and monopolistic competition. *The Review of Economic Studies*, 43(2):217–235, 1976.

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

T. Tabuchi. Urban agglomeration, capital augmenting technology, and labor market equilibrium. *Journal of Urban Economics*, 20(2):211–228, 1986.