The Economics of Agglomeration and Firm Location

This article reviews, summarizes and critiques the existing literature on agglomeration and firm location decisions. The strengths and weaknesses of agglomeration economies as an explanation of current settlement trends are discussed. The article covers methods of measuring agglomeration economies, including surveys, indices, production function models, and simulations. Most studies conclude that agglomerative forces are strong. There is a need to refocus research on the issue of whether current subsidies provided to firms moving to the urban fringe are needed, and how the benefits of agglomeration are distributed.

Introduction

If firms were animals, a behavioral scientist trying to describe their traits would observe that firms tend to be found in herds and usually migrate toward the largest watering holes. Industry and population statistics indicate location choices of business are predominantly urban. In the 1980s, metropolitan employment in the US grew twice as fast as non-metropolitan areas (Barkley 1993). This follows a long trend of migration from rural areas over several decades. From 1950 to 1990, the percentage of population located in or around metropolitan areas has grown steadily from 56 to 77 percent (Statistical Abstract of the United States 1990:910). Heavy industry and population concentrations in urban areas ensures that the majority of present and future job growth is centered around cities.

A second trend is the movement of firms to the urban outskirts. Relatively undeveloped rural areas adjacent to cities are rapidly changing as former farm land is replaced with office complexes. The rise of suburbs as important centers of economic activity has been substantial; major corporate headquarters are increasingly located in the suburbs. Chicago, Los Angeles and New York now have equal numbers of suburban and downtown corporations (Stanback 1991). Only businesses highly dependent on information-intensive, one-on-one interactions still find significant locational advantage in downtown areas (Maki 1992). The suburbs of cities are emerging as the places where most firms seem to find their greatest locational advantage.

These trends are curious because many of the amenities important to business are no longer exclusive to cities. Advancements in transportation and telecommunications have reduced the unreliability and costs of doing business over long distances. In contrast, the costs of city life and work include greater fixed expenses for business, such as higher land rents, greater energy costs, stricter environ-
mental and air quality regulation, and more substantial labor costs. Urban residents drawn to the economic opportunities of cities face more crime, pollution, congestion, higher living expenses and increasingly longer commutes. Social costs include the loss of scenic rural country to suburban sprawl, environmental quality concerns and increasing disparities in per capita income for nonmetropolitan versus metropolitan residents. Despite these and other costs, cities and their surrounding suburbs continue to be business hubs, and there is an increasing sense that major metropolitan areas are the cradles of big ideas, new technology, and innovation.

Given these costs, firm location decisions are undoubtedly a critical concern for both policy and research agendas. Questions of importance are:

1) Do the urban preferences of firms ignore gross negative externalities, which, if internalized, would alter present settlement patterns?

2) Are the disadvantages of nonmetropolitan sites so prominent that state efforts to lure economic development to rural areas are relatively ineffectual?

3) Are the benefits of firm clustering so great that policy efforts aimed at redirecting industry location compromise the competitiveness and profitability of industry?

4) Is there such a thing as an optimal city size, and have major metropolitan areas exceeded that ideal?

Key to answering these questions is a better understanding of the incentive structure driving firm location. Acknowledging that uneven economic growth is in part due to the location choices of firms, economic research has attempted a rigorous approach to the central question of why industry so consistently clusters in and around metro-politan areas. If it is assumed that a firm rationally selects its site, the choice of an urban location signals that there are cost advantages associated with the location. While no single theory can satisfactorily address the nuances involved in all firm location decisions, the most consistent theoretical explanation used to describe the benefits accruing to firm location choices is agglomeration economies.

The purpose of this paper is to provide a review and assessment of the literature, mainly from the discipline of economics, that relates to how agglomeration externalities affect firm location decisions. The next section provides a working definition of agglomeration economics so the reader can appreciate the nuances of some of the studies reviewed later in the paper. Then some factors that can lead firms to agglomerate without increased efficiency are discussed. Next basic methods of assessing the degree of agglomeration economies are presented, and then evidence on diseconomies associated with agglomeration is discussed. A concluding section sets out new research directions that could have more direct policy implications than the research to date.

**Agglomeration Economies:**

**A Working Definition**

Weber defines agglomeration as "an "advantage" or cheapening of production or marketing which results from the fact that production is carried on at some considerable extent at one place..." (1929:126). In an economist's vernacular, there are four central components of agglomeration benefits: **transfer economies, internal economies of scale, localization economies** and **urbanization economies** (Nourse 1968).

Transfer economies refer to the transport savings a firm reaps by locating close to its
markets or near other firms that supply it with inputs. These cost reductions are positively related to the bulk of a firm’s inputs and outputs. Savings on information are also a form of transfer economies and are dependent on the complexity of information that is needed to sell consumers a product or receive an input from a supplier. For example, the need for face-to-face contact may increase the advantages of locating close to clients and near input suppliers. This is likely particularly true for the service industry, which produces goods that are often tailored to individual customer demands.

Internal economies of scale are reductions in the long-run average cost curve of a firm that occur when output expands. They accrue as the firm itself increases its scale of production. Clustering can positively influence a firm’s internal unit production costs. For example, if firm B locates in firm A’s vicinity and requires firm A’s output for its own production process, firm A is able to expand its market and decrease average costs per unit. In general, internal economies of scale can be brought about by a variety of factors including division of labor, integration of processes, economies of increased dimensions, economies of specialization, technology, learning, or improved organization of production (Meyer 1977).

Localization economies are cost savings that are internal to the industry but external to the firm. They occur when many firms in the same industry locate in an area to capture the external benefits of clustering. The source of the agglomerative benefit is in the size of the industry, rather than the size of the firm. The industry reduces its long-run average cost curve by increasing its output. Localization economies describe the benefits that accrue to firms that locate together, regardless of whether they choose an urban site. A commonly cited reason for concentration of industry in one locale is the development of a sizable and specialized labor force. Other benefits include waste processing, research and development, and raw material accessibility (Nourse 1968).

Some elements of how localization economies function have sparked debate in the literature. Porter (1990) hypothesizes that proximity creates competition among a ‘diamond’ of firms, factors, demand and institutions. This competition induces investments in factors of production, thereby creating efficiencies and demand, which in turn create internationally competitive firms and nations. A number of authors suggest that it is not competition, but better collaboration and networking that are the driving forces in localization economies. Saxenian (1994) posits that the size of localization economies may depend on the industry’s culture; industries with a high level of trust and collaboration across firms will be in a better position to produce the gains associated with localization. Proximity makes it easier for manufacturers in an industry to get experience dealing with one another, which leads to trust, fostering collaboration and higher regional growth; however once trust is established it cannot always be maintained (Harrison 1992). Gertler (1992, 1993) points out that proximity among producers and buyers facilitates communication and thereby iterative development of goods and technologies more appropriate to local culture and production systems. Other authors suggest that networks are more important than competition, but that networks are not spatially bounded (D’Cruz and Rugman 1992, 1993; Rugman and D’Cruz 1993).

Localization economies also exist in retail and other output markets. Often retailers agglomerate despite their rivalry. Goods that are not perfect substitutes are often best assessed by inspection, explaining why chain dealerships, chain restaurants, and stores selling similar products tend to cluster (Hoover and Giarratani 1984).

The broadest form of agglomeration benefits are urbanization economies. These are the savings external to both the firm and industry.
They develop because business activity of all types is concentrated in metropolitan areas. The city itself fosters cost savings for the firms clustering in it. These benefits of urbanization come in many forms, and it is this fact that makes measurement of agglomeration economies difficult. Examples of possible cost savings due to an urban location include greater access to financial services, capital, infrastructure, labor and the need for smaller inventory due to close wholesale sources. Specialization is possible and products and services are available to purchasers at a lower cost than without urbanization economies (Hoover and Giarratani 1984). Economies of scale also exist in the publicly provided infrastructure needed to support business operations (Isard 1956). The existence of strong urban economies does not preclude the possibility of significant external costs that may not be borne by the firm when it chooses an urban site. A classic negative externality of this type is pollution.

ALTERNATIVE EXPLANATIONS OF FIRM LOCATION

All firm location decisions cannot be lumped under the rubric of agglomeration. The significant attention that agglomeration economies has held in explaining site location has prompted criticism that it is a conceptual ‘repository for the untidy aspects of location theory,’ (Bopp and Gordon 1977:125). There are forces beyond agglomeration which influence site selection; several are mentioned here.

Search Costs

Pascal and McCall (1980) accept agglomeration as a response to the advantages of clustering, but contend that agglomeration economies cannot explain the full force of urban location decisions. The authors turn to the importance of risk in determining the location for a firm entering an industry. The assumption of perfect information in many economic models implies that new firms know the best place to locate business. Given the realities of imperfect information, new firms will try to lower their search costs by observing and then imitating the location choices of successful businesses. In doing so, they seek to increase the probability of their own profitability. While ‘monkey see, monkey do’ tendencies may predominate when a young, risk averse firm faces uncertainty, search costs cannot address how older establishments originally selected the particular sites that younger firms eagerly copy.

Current Business Trends

There is an undoubted evolution of North American industry from manufacturing toward services. Service industries are freer to locate near their markets than are other types of industry, particularly those heavily dependent on natural resources or manu labor. Service goods by definition target concentrations of population. The increased migration of firms to urban settings may simply be part of the rising prominence of services in the national economy. Once a concentration of service firms is established in a locality, firms relying on external producer service providers may find they save search time and information costs by locating close to the service clusters. Other business trends also influence firm location. For example, there is a growing trend for corporations to physically separate production, management and marketing functions of business. Given the revolution in telecommunications and computer technology, modern firms can dismantle activities, locating operations where each division can take advantage of different amenities. Routinized activities can be readily placed in low-rent regions while information-intense operations are based in urban locales.
Noneconomic Considerations

Economic theory tends to focus strictly on the cost-saving factors bearing on a firm's decision to choose a particular site. In practice, a firm's choice of a site may not be rigorously scientific nor reflect the firm's recognition of and search for a site holding agglomeration benefits. Blair and Premus (1987) stress that the location decisions of firms cannot be viewed as an isolated activity but as part of an intricate, overall corporate strategy. Building an additional facility is only one of the many options available to existing businesses facing capacity constraints (Schmenner 1979). If another site is determined to be necessary, the firm typically assembles a site selection team that chooses a new location by considering a myriad of factors (Blair and Premus 1987). Economic criteria include a firm's desire to enter new markets, corporate interest in integrating or segregating its activities, and management concern over company visibility.

Noneconomic factors are also important. One survey examining choices for corporate headquarters found that CEOs have substantial opportunity to exert their personal preference in selecting a locale (Blair and Premus 1987). In a mail survey reviewed by Blair and Premus (1987), 41 percent of respondents indicated personal concerns to management were a critical consideration in location decisions. For existing business, these concerns were important for 29 percent of survey respondents.1

New businesses tend to locate near the entrepreneur's home, not necessarily where the enterprise has the best opportunity for success. Krugman (1991) posits that noneconomic factors create initial activity in an area, and then cumulative causation gives rise to industrial concentration as firms and workers both prefer larger markets. The role of subjective criteria in location choices underlines the importance of not mistaking location choice as being a simple function of agglomerative forces.

SURVEYS AND INDICES

Acknowledging the importance of factors beyond agglomeration economies assists in understanding the totality of issues that firms consider in choosing a location. The existence of these forces does not detract from the importance of developing an economic theory of firm location that addresses agglomeration economies. Economists use several methods to quantify agglomerative forces. Surveys and indices, among the simplest of these methods, are discussed here. Other more mathematically complex methods are discussed in succeeding sections.

Surveys

The most empirically straightforward way to measure the importance of agglomeration factors in firm location is to directly survey businesses that have started up, expanded or relocated. While this approach is unencumbered by complex methodology, the existence of agglomeration benefits is difficult to assess via survey because it is obviously not possible to ask management, 'Did you choose the site to capture the positive external benefits brought on by agglomeration?' Proxies that are generally used to identify agglomeration forces include questions about neighboring suppliers of inputs, area labor availability and proximity to markets.

Surveys show location decisions of US firms in the 1940s and 1950s as dominated by access to markets, labor and raw materials. These variables are less critical now (Blair and Premus 1987). Fortune magazine's review of the location decisions of 1000 of the largest US industrial firms found market access to be the most important factor in expanding. When management was asked to consider factors influencing previous plant locations, they
identified labor availability as a chief consideration, an indication that agglomeration economies may play a more important role in the future than they have in the past. Survey research also reveals that the important elements of location decisions vary across firm type. High-tech executives ranked transport of inputs and market access low in site selection but considered local availability of specialized labor and distance to universities as the first and third most important attributes (Blair and Premus 1987).

In general, accurately constructing a survey to use as a data set for testing location hypotheses has been problematic (Dorf and Emerson, 1978). Barkley and McNamara (1994) analyzed the survey process for inconsistencies between survey responses and actual location decisions taken. They found consistent information for easily observed criteria such as highway proximity. For other characteristics, they found consistencies when the data were disaggregated by plant size and labor skills requirements. They also found that while non-economic characteristics may not attract firms, these factors may dissuade business from considering a given community.

Indices

Data on number of firms, population, and markets have been used in a variety of ways to construct indices that determine where agglomeration forces are the strongest. For example, Wheaton and Shishido (1981) developed a simple agglomeration index for use in an economic model:

\[
H = \sum_{i=1}^{n} \frac{P_{i}^{2}}{P^{2}}
\]

(1)

where:

\[
P_{i} = \text{Number of firms in industry } i
\]

\[
P = \text{Total number of industries}^2
\]

This specification can be used to measure agglomeration if it is assumed that a heavy concentration of industries is equivalent to industries being tightly interwoven (i.e. strong forward and backward linkages) so locating together helps capture the external benefits of agglomeration.

Dieterink and Nijkamp (1988) developed a more sophisticated index to analyze the effects of agglomeration economies on innovation and research and development. Their index is a function of city size, distance to the main city center and distance to smaller, neighboring towns.

Marcus (1965) uses growth rate differentials to index agglomeration benefits. The model computes expected growth of a regional industry as follows: the ratio of the growth rate of an industry in some region relative to the growth in population of the state is expected to equal the ratio of national industry growth to national population growth. If the industry grows at this expected rate, then the difference between the growth rates will be zero. If this difference is nonzero, then this residual is assumed to be caused by agglomeration effects.

Other indices incorporate the linkages between industries and are careful to distinguish between industrial concentration and true economic linkages between the industries that may produce agglomeration benefits. The central hypothesis is that if agglomeration economies are important in the firm’s location choice, then pairs of industries linked by flows of goods and services will tend to site close together. No agglomeration benefits are assumed if linked firms do not choose to operate in the same area. To determine which industries tend to cluster together, correlation coefficients or factor analysis is used. Results provide a gauge of spatial association. However, correlations express geographical ten-
dencies rather than functional association; consistent spatial associations do not imply that there are economic linkages between sectors. The location tendencies of unrelated industries locating in the same area must be separated from those firms that locate together because they are inter-related and experience agglomeration benefits. This has most typically been achieved by input-output analysis. Examples of this approach can be found in Bergsman et al. (1972), Lever (1972), Streit (1969), Bopp and Gordon (1977), and Dorf and Emerson (1978).  

Gravity models have also been a common approach to modeling the movement of goods and services across space. Gravity models assert an inverse relationship between movement and distance (Zipf 1949). Interaction and trade can be explained by the length and intensity of movement required, and firms will try to minimize their effort expended. Taylor (1973) adapted this idea to analyze regional agglomeration economies. The author attributes differences between expected and observed trading patterns to 'the immediately local industrial environment' (p. 393) that offers benefits to firms locating there.

Production Function Models

Indices attempt to measure agglomeration by exploring economic linkages across sectors. This approach has limitations. Meyer (1977) warns that linkages may not be a good gauge of agglomeration benefits because linked growth between industries may be due to factors having nothing to do with capturing external economies. More rigorous work has used production function models to directly measure the external benefits of localization and urbanization. This section introduces the basic economic theory of production functions, explains how these models are used to measure agglomeration and summarizes research results using these methods.

BASICS OF PRODUCTION FUNCTION MODELS

Neoclassical economic theory asserts that the firm chooses its input levels and supplies output to the market to maximize its profits. Profit maximization is limited, or constrained, by the technology available to the firm. A production function is a mathematical equation describing how the firm's inputs are combined to produce its output. In its most general form a production function is expressed as follows:

\[ Y = g(\cdot) f(K, L) \]  

(2)

where:

- \( Y \) = level of output
- \( g(\cdot) \) = scale effect parameter, a constant
- \( K \) = capital
- \( L \) = labor

The theoretical model above asserts that all inputs can be described as a form of labor or capital. In equation (2) output is not only dependent on inputs, but also on \( g(\cdot) \), the scale parameter. The scale parameter is estimated econometrically and describes the impact that all other factors (beyond labor and capital) have on production. The scale effect is exogenous—it is faced, rather than chosen by the firm. It can be thought of as a magnification term, some coefficient that is independent of input usage but affects overall production.

THE COBB-DOUGLAS MODEL APPLIED TO AGGLOMERATION RESEARCH

For empirical estimation, it is necessary to choose a functional form for the theoretical model described in (2). The selection of a specific equation implies making an assumption about the technology available to the firm.
A common specification is the Cobb-Douglas function:

$$Y = e^{\lambda} K^\alpha L^\beta \quad \alpha, \beta > 0 \quad \lambda > 1$$

(3)

where:

- $Y$ = level of output
- $\alpha, \beta$ = returns to scale parameters
- $e^\lambda$ = scale effect parameter; technical change exponent
- $K$ = capital
- $L$ = labor

The Cobb-Douglas model is the simplest production function model used to measure the external benefits of firm location. Two other models are actually more prevalent in the literature, the constant elasticity of substitution (CES) function and the Dhrymes specification. The CES function is a more general form of the Cobb-Douglas model, while the Dhrymes (1965) specification starts with a CES model and uses an assumption—perfect competition in input markets—to eliminate the need for capital data in the model. There are two basic approaches used to measure agglomeration economies with production function models (Duffy 1987): estimating the scale effect parameter and estimating the returns to scale parameters. Each of these methods is described here.

The scale effect of the production function

(e.g. Segal 1976; Sveikauskas 1975; Moomaw 1985; Greytak and Blackley 1985). This approach examines differences in the scale parameter across space. Each city or geographic area under observation is assumed to have an aggregate production function of the same form (for example, Cobb-Douglas). That is, it assumes inputs for identical industries in different cities are used in more or less the same combinations and that all firms have access to the same technology. Thus, differences in production output across cities stem from differences in this scale effect. The higher the scale parameter in value, the more productive economic activity is in one area over another. For agglomeration research, the scale effect, $g(\cdot)$, is specified to be a function of some variable that expresses agglomerative behavior. Population is assumed to be a reasonable proxy for agglomeration benefits, and often the scale effect is specified to be a function of population.

Important to this technique is whether one assumes technical change affects all factors of production equally. If technical change impacts all inputs in production equally, then it has a property called Hicks-neutral productivity. The Cobb-Douglas specification assumes technology is Hicks-neutral; a technical change altering $\lambda$ changes the average factor productivity of the firm but the marginal rate of substitution between inputs is unchanged. Most attempts to fit production functions to output data for cities assume Hicks-neutrality. This has been questioned by some who argue that factor bias is more likely and that assuming Hicks-neutral technology is a mis-specification of the model, thus biasing scale parameter estimates across economic areas (Carlino 1979).

Although this method does not reveal why the scale parameter values may differ, it does measure agglomeration's role in producing more output for a given level of input. Because the exercise assumes that inputs and technology are constant across space, possible differences in the marginal products of labor and capital are not captured using this approach.

In the returns to scale parameter

(Shefer 1973; Carlino 1979; 1982). Returns to scale (RTS) parameter estimates give economists an idea of how much production
changes when all unit inputs are changed. In the Cobb-Douglas model, RTS can be found by summing the exponents attached to labor and capital. If their sum equals 1, constant returns to scale (CRS) exists. That is, doubling inputs will double output. These values are used to measure returns to agglomeration. For example, if an industry in a big city has higher RTS than the same industry in small cities, it is assumed that the big city allows for greater productivity given the same factor usage.

**Empirical Results**

A number of empirical studies using the production function approach are found in the literature. How do the results compare? Since each study tends to use different data, looks at a different region and time period, and typically uses a unique estimation procedure, one might expect wide differences in results. In fact there is broad agreement—most studies (Segal 1976; Moomaw 1981; Sveikauskas, 1975; Gretyak and Blackley 1985; Tabuchi 1985; Calem and Carlino 1991) find increasing factor returns with city size, although the estimated rate of increase varies. Shefer (1973) finds that returns to city size, while positive, vary from one industry to the next. A minority, Carlino (1979) and Moomaw (1986) is only slightly less optimistic—finding increasing city size has positive impacts on some industries but negative impacts on others. Duffy’s (1987) results indicate a sinusoidal relationship between returns to scale and city size. Those studies that attempt to measure optimal city size find it to be in the millions—2 million for Calem and Carlino (1991), 5.5 million for Kawashima (1975). Nakamura (1985) finds that optimal city size depends on the industry, with medium-sized Japanese cities best for manufacturing. Carlino (1982) finds the optimal size to be 3.3 million but declining over time. Approaching the problem from a different perspective, Hansen (1990) finds declining productivity with increasing distance from the central business district. The weight of these studies is that agglomeration economies are important, and that most localities can become more efficient by growing.

**Limitations of the Production Function Approach**

Production function estimation probably provides a more precise measurement of agglomeration than does a simple index. But production function estimates have some problems that may affect the overall accuracy of the method.

**Data Problems**

Most studies lack firm-level data and resort to using the industry as the unit of observation, often employing two-digit SIC data, which can lead to upwardly biased returns to scale estimates (Gretyak and Blackley 1985; Peterson 1990). One-factor models, such as the Dhrymes specification, lead to heteroskedasticity and multicollinearity.

**Capital valuation**

Capital goods cannot be summed directly and differ in type and value, so capital value estimates are often substituted. This is usually investment minus depreciation and leads to changing capital values without a corresponding change in capital stock, or vice versa. For example, most depreciation schedules will make the value of investment in a building little or nothing after thirty years, even if the building is still in use.

**Time series vs. Cross-Sectional Analysis**

Some researchers prefer time-series analysis because it eliminates the possible effects of locational attributes such as ports.
sectional analysis has the advantage of controlling for technical change, but a disadvantage in that the production function and input prices are heroically assumed to be identical across cities.

Hicks-Neutrality

Using a Cobb-Douglas production function implies that Hicks-neutrality best describes the innovation process. This approach does not allow for differing capital-to-labor ratios across city size (Greytak and Blackley 1985).

MATHEMATICAL PROOFS
AND SIMULATIONS

This type of literature does not attempt to measure agglomeration economies; no empirical tests are performed. Much of this more theoretical work employs agglomeration as an input for explaining the size of cities and other urban phenomena. An example of this approach is Helsey and Strange (1991). They analyze the role of risk in capital markets as a source of positive agglomeration economies. They contend that investors face risk in providing credit to investment projects and that the probability of a default drives the owner of capital to consider the next best option for his capital should the initial investment fail. Agglomeration economies exist because the salvage value of fixed capital assets is higher in a more diversified economy. Other theoretical work dealing with agglomeration and firm behavior includes Fujita (1985a, 1985b); Grimaud (1989); Ogawa and Fujita (1980); and Papageorgiou and Thisse (1985).

THE ROLE OF DISECONOMIES
IN AGGLOMERATION

The analysis and measurement of agglomeration economies through a variety of modelling and empirical tools supports the conclusion that benefits accrue to firms locating together, most often in urban areas. Observed migration patterns and the successful measurement of increased efficiency of cities provide evidence that the agglomeration pull is much stronger than the disadvantages attendant to an urban site.

Research has only begun to address the external diseconomies that exist in the face of agglomeration benefits yet are not incorporated as a cost in the firm's location decision. While modelling has provided a rough gauge of agglomeration savings, Carlino (1979) and others correctly point out that these efforts provide a rough approximation of net benefits that accrue to firms locating in a metropolitan area. Clear delineation of external diseconomies and agglomeration economies is needed before firm settlement patterns can be deemed socially optimal.

These external costs are less understood and modelled only occasionally. Westaway (1974) discusses several costs frequently ignored in a social accounting of location decisions, among them the 'way in which regional variation in the occupational distribution of labor affects social mobility' (p. 152). Edel (1972) also examined the distribution of gains and losses due to city growth and found them to be regressive. He also attempted to measure the diseconomies of congestion through decreases in land values. His findings indicate that on average, agglomeration economies surpass diseconomies for cities up to 0.5 million in population but that for large cities it is unclear which force is dominant. Because the private costs and benefits of firm location are not necessarily those borne or received by society, policy based on agglomeration findings must recognize possible differences in private versus social costs.

SUMMARY AND CONCLUSIONS

Agglomeration economies are instrumental in explaining firm location. The importance of examining net social costs and benefits of firm

location remains an important research task. Also significant is the need to isolate the importance of localization versus urbanization economies. These two categories of agglomeration benefits are strikingly dissimilar and their relative magnitudes in firm location dictate different policy directions. If the bulk of agglomeration economies a firm experiences come from the intrinsic savings of a city location, efforts to lure industry to rural areas will likely require a substantial level of subsidy to compensate for this loss. But if firms receive benefits not so much from a metropolitan site but because these areas are home to other linked industries (e.g. localization economies are more significant that urbanization economies), transporting interdependent firms *en masse* from a city to a rural location would not, in theory, adversely affect their cost structure.

To date, efforts to move industry out of densely populated areas via state or federally sponsored incentive programs have enjoyed only modest success (Schmid 1991). Returning to the fundamental economic reasons why firms locate in and around metropolitan areas may offer suggestions about the way in which a more socially optimal population distribution can be attained.

Agglomeration research has promise in answering these questions but has tended to go in directions that are analytically interesting but not very useful in supporting policy direction. We have many studies telling us optimal city size is in the millions--more work in this area is probably not needed. What we don't know is how the benefits of efficiencies gained through agglomerative tendencies are distributed. Are certain categories of firms or individuals hurt by agglomeration, or does everyone benefit from current agglomerative trends? Another important policy question, un-answered in the literature, is whether the current policies of subsidizing infrastructure development are needed--if firms are capturing a large portion of the gains associated with agglomeration, then they should not need extra incentives to move to the urban fringe. Redirecting efforts in this area should yield work that is important in helping scientists define a research agenda.

Notes

1. The mail surveys tended to focus on smaller firms, thus possibly exaggerating the importance of personal preferences in the business decision-making process.
2. The index has been modified slightly from the author's original presentation which focuses on population concentration.
3. Two industries or firms can be strongly correlated and have nothing to do with one another. For example, export shipping and the fisheries industry both tend to be located in coastal areas. But neither industry strongly relies on the other for inputs of production nor shares a common output market.
4. See Selting, Allanach and Loveridge (1994) for a more detailed discussion of these specifications.
5. See Selting, Allanach and Loveridge (1994) for a more detailed discussion of these specifications.
6. Since units matter, the researcher must be careful to ensure that the unit of measurement is the same across industries and regions to make comparisons valid.
7. For discussion and examples of estimation procedures, see Selting, Allanach and Loveridge (1994).
8. Other properties of interest include a factor's share of productivity, technical change and homogeneity properties. See Eatwell, Milgate and Newman (1987) for a discussion of these topics.
9. Similarly, if their sum is greater than 1, increasing returns to scale exist (IRTS) and doubling inputs will more than double output. If their sum is less than 1, decreasing returns to scale exist (DRTS) and doubling input will less that double output.
10. For a more detailed review of methods and results of the studies summarized in this section, see Selting, Allanach, and Loveridge (1994).
11. Note the similarity of this argument and labor market theories presented earlier.

Acknowledgements

This work received support from the North Central Regional Center for Rural Development, the University of Minnesota, and West Virginia University. The authors thank Wilbur Maki, George W. Morse, Nick Piggott and two *Great Lakes Geographer* referees for comments on earlier versions of this work.


Moomaw, R. L., 1986. 'Have Changes in Localization Economies Been Responsible for Declining Productivity Advan-