Geographic Concentration and Establishment Size: Analysis in an Alternative Economic Geography Model*

Thomas J. Holmes and John J. Stevens

February 2002

Abstract

Big cities specialize in services rather than manufacturing. Big-city establishments in services are larger than the national average while those in manufacturing are smaller. This paper proposes an explanation of these and other facts. The theory is developed in an economic geography model that is an alternative to the standard Dixit-Stiglitz structure. In our tractable structure that has potentially wider application, firms have monopoly power in local markets, but are price takers in export markets.

Keywords: geographic concentration, establishment size, transportation costs, new economic geography

JEL Classification: L11, F10, R10, R30

*Holmes, University of Minnesota and Federal Reserve Bank of Minneapolis (holmes@econ.umn.edu). Stevens, Board of Governors of the Federal Reserve System (John.J.Stevens@frb.gov). Holmes acknowledges support from the NSF through grant SES 9906087. Any views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

The new economic geography has had a substantial impact on our understanding of the specialization patterns of geographic areas. So far it has not found much application in helping us understand plant-level observations. This is a surprise, since at its core, this theory is about plant-level scale economies. We can expect that this theory would have much to say about regional variation in plant size. For example, this theory might help account for the fact that plants are substantially larger in locations that specialize in an industry, as compared to other plants in the same industry (Holmes and Stevens (forthcoming)). This potential is in sharp contrast with other spatial theories, such as those emphasizing knowledge spillovers and externalities, which arguably have no content for plant-level phenomena.

This paper aims to use the new economic geography ideas to begin to integrate the study of regional specialization and the determinants of plant size. To explain the contribution, we divide the remainder of the introduction into two parts. The first part lays out a number of facts about cities and establishment size. It shows that there is a strong relationship between specialization patterns of cities and establishment size. It then proposes a parsimonious explanation that can simultaneously account for all of the facts. The second part provides an overview of the theoretical model that formalizes our explanation. While the model follows the spirit of the new economic geography literature, it departs from this literature by not using the standard Dixit-Stiglitz structure. The second part of the introduction motivates our choice of this alternative structure to make our point. It also argues that our alternative structure may have broader applicability.

1.1 Facts and an Explanation

It is well known that cities of different sizes specialize in different types of economic activity. Table 1 documents specialization patterns for U.S. cities using data from the 1997 Economic Census. The 328 primary metropolitan statistical areas (PMSAs) are divided into small cities (PMSA population under half million), medium (half million to two million), and large (over two million). The table reports location quotients (LQ) for major sectors for the three size classes. The LQ equals the share of industry sales occurring in establishments located in the given city size class, divided by the share of population in that size class.[1] The table shows that service-type activities, such as wholesale trade and finance, are heavily concentrated in the large cities (the LQ increases from approximately .5 in the smallest cities to approximately 1.3 in the largest cities for the three service-type sectors). Manufacturing activity is concentrated in small cities (the LQ declines from 1.15 to .86). Retail is spread out evenly across city size types (the LQ is approximately one in each case). These specialization patterns are well known.[2]

Less well known is the link between specialization patterns and establishment size. Table 1 also reports *Size Quotients (SQ)*. As in Holmes and Stevens (forthcoming) we define a size quotient in sector i and location j to be the ratio of mean establishment size in sector i at location j to the mean size in sector i in the U.S. (We use sales as our measure of size, just as we use sales to calculate the LQ). For example, the manufacturing SQ for small cities is 1.25. This means that average manufacturing plant size in small cities is 25 percent larger than the U.S. average size. The clear pattern of this table is that establishment size declines with city size for manufacturing, but increases with city size for the service-type sectors. Retail establishment size also increases in city size, but at a much less pronounced rate. In Holmes and Stevens (forthcoming), we further document the strong connection between specialization and establishment size.

We begin our explanation of this set of facts by invoking the fundamental elements of the economic geography literature: transportation costs and scale economies. Transportation costs are particularly high in the service sector. Face-to-face communication is often crucial with services, and the transportation costs of moving people surely dwarf the cost of moving

[1] Major sectors are defined to include the NAICS major sectors with more than $500 billion in receipts. For the Finance and Insurance sector (NAICS 52), payroll is used instead of sales because PMSA-level sales for sector 52 is not released. In the Wholesale Trade sector and the Finance and Endurance sector, the Census withheld data for a few PMSAs. But the coverage of disclosed data was greater than or equal to 95 percent of all establishments for all the reported figures.

[2] The patterns regarding services and retail are obvious and long-standing. The pattern involving manufacturing is a more recent phenomenon. Early in the twentieth century, large cities were centers of manufacturing. However, by the end of the century manufacturing had shifted to smaller cities. Mills and Hamilton (1994) note a "dramatic shift of manufacturing from large to smaller MSAs" that occurred over this time period.

goods. Service sector activity concentrates in large cities because large home markets make it possible to both economize on the cost of moving people and to achieve economies of scale. The manufacturing sector concentrates in small cities. With the relatively low cost of shipping goods rather than people, manufacturing plants in small cities can obtain scale economies by shipping to a national market. The retail sector is spread out evenly across city sizes because transportation costs are prohibitively high and retail services are not traded.

To these basic ideas from the new economic geography literature, we introduce a small perturbation. We observe that in any sector, or even any broadly defined industry, there are likely to be some subsets of products within these industries that are difficult, if not impossible, to trade, even if most of products in the industry are capable of being traded. In most any industry, some goods must be custom-made rather than standardized, and it is often convenient to have the custom-made articles produced near the point at which they will be consumed. This can be true for clothing, furniture, printing, and so on. In addition, in many industries it is important to have the final stage of production occur near the final point of use as some products may be difficult to ship after final assembly. Our explanation for why manufacturing plants are small in large cities is simply that such plants tend to only engage in custom work, whereas plants in small cities are large because such plants produce the more standardized products that are exported to other cities. Analogously, service sector establishments are small in small cities because they tend to only do the custom work that cannot be traded. The retail sector is different from the manufacturing and service sectors because all retail establishments only sell to local markets. Retail establishments are larger in bigger cities because of the larger local market base, but the size difference between large city and small city establishments is not as great as it is for services. Unlike the retail sector, large city establishments in the service sector are exporters, and this accentuates the relative size difference between establishments in large and small cities.

Our simple explanation for the facts is consistent with an additional fact documented by Bernard and Jensen (1995, 1999). They compare manufacturing plants that export (to other nations, not cities) with plants that do not export and determine that exporters are substantially larger than non-exporters. One explanation for this fact in the trade literature is that there may be a *fixed cost* component to exporting (see Melitz (1999)). Our explanation is simply that the entire reason for existence of these small manufacturing plants is to provide custom activities that cannot be traded.

1.2 Our Formalization

We formalize this story in a model that is something of a radical departure from the standard Dixit-Stiglitz structure that is the basis for the new economic geography literature. In our model there is a continuum of locations. At each location there is at most a single producer of any one good, so such a producer has a monopoly in the local market. In export markets, however, the firm competes with "foreign" producers (i.e., from other cities) of the identical product, so export markets are perfectly competitive. In this model of monopolistic competition, a given firm has a monopoly in one market (the local market) and is competitive in a second market (the export market). There is also free entry.

In our parsimonious model, all products in the economy have an identical production function that includes a region of increasing returns to scale. Products only differ in transportation cost. There are products with generally low transportation costs (we call them manufactured goods), products with generally high transportation costs (services), and products with infinite transportation costs (retail). In addition, for both manufacturing and services, we assume there is some component of demand that can only be satisfied by local production (i.e., some component that must be custom-made). Locations (cities) differ in population size. The equilibrium pattern of specialization in this economy and the size distribution of plants follows the story outlined previously.

We were led to depart from the standard Dixit-Stiglitz structure as it is ill-suited for analyzing establishment size. As is well known, all of the action in the Dixit-Stiglitz model is on the extensive margin—the addition or substraction of new firms (See Holmes (1999) for a further discussion of this point). The equilibrium size of firms is fixed; it is, in essence, a parameter of the model. Our framework allows for a relatively general cost structure and for the possibility of potentially rich variations in firm size.

In addition to its suitability for our purpose, we believe that our structure may be of broader interest and we are currently employing a variant of this model in another application (Holmes and Stevens (2002)). At the margin the firms in our model behave competitively, since they are competitive in export markets. As a result, the definition and construction of equilibrium in our model has properties that are similar in many ways to standard models of perfect competition. This is a useful feature because models of perfect competition tend to be much more tractable and allow for greater generality than do standard Dixit-Stiglitz models. Our structure can potentially be justified on empirical grounds. In our model, the demand elasticity for local sales is strictly less than for export sales (the latter of which is infinity). It seems empirically plausible to us that for most firms, the elasticity of local demand would be less than that for exports. This is certainly consistent with anecdotal evidence that firms often set lower margins in export markets than in local markets (these situations often lead to complaints about "dumping"). In the standard Dixit-Stiglitz model, a crucial element of the framework is that the elasticity be the same in both the local and the export markets.

A basic result of our model, that the medium transportation cost sector concentrates in the big cities, while the low transportation cost sector concentrates in the small cities, is a well-established central result in the literature. This is a main point in the original Krugman papers (i.e., Krugman (1980), Krugman (1991)). Our result is most closely related to a result in Amiti (1998), where transportation cost varies across industries but everything else is held fixed. (This contrasts with the original Krugman analysis where the low transportation cost sector has a different technology (constant returns) and different preferences (no product differentiation) as compared to the high transportation cost sector.) Our analysis is different from Amiti in that we introduce non-traded components of demand in all sectors and we are concerned with the scale of establishments. Beyond that, the differences are that we have a continuum of locations rather than two locations, and we use our new structure while Amiti works in the standard Dixit-Stiglitz structure. An appealing feature of our continuum analysis is the way it highlights the intuition for our specialization result. We derive for each sector an object that is analogous to a bid-rent function from urban economics. These bid-rent functions pin down the specialization patterns across locations in the same way that they pin down land use patterns in urban economic models.

2 Model

We describe a model with a continuum of locations that vary in population. We simplify by assuming that the population of each location is fixed and therefore ignore the issue of labor mobility. Thus in its formal structure, our model is more akin to a trade model like Krugman (1980) rather than a regional model like Krugman (1991) with factor mobility. In applying our theory to help us understand data across cities, we of course recognize there is labor mobility across cities. We view our model as a first cut at determining equilibrium with a continuum of locations. We believe that if we were to extend our model to endogenize city size, the forces at work contributing to city specialization that we highlight would remain.

Since the key assumptions of the model have to do with goods and their transportation costs, we first provide some motivation for the commodity structure used in the model. There is a continuum of goods. At each location, there is a certain component of the local demand for the good that must be satisfied by local production and a second component that may be satisfied by local production or imports. Consider, for example, a service good called "grocery wholesaling". One component of the demand for this good is for the wholesaling of milk. For this component it is important that service be provided locally because of perishability. A second component is the wholesaling of crackers, and this component could be satisfied by imports since issues of perishability do not arise. A second example is cabinet manufacturing. One component of demand is for cabinets that are built into the wall. For custom work such as this, it is helpful to build the cabinet near the place it will be installed to better coordinate production (the model assumes this is a necessity). A second component of demand is for free-standing cabinets that can possibly be imported from other locations (though with some transportation cost). The model assumes that while these two components are different on the demand side, they are the same on the production side; a grocery wholesaler can provide cracker wholesaling and milk wholesaling as part of the same production process, likewise a cabinet maker can produce built-in and free-standing cabinets. The results would be qualitatively the same if there were assumed to be different production processes. Let us now turn to a formal development of the model.

2.1 Locations and commodities

There is a continuum of locations indexed by $i \in [0,1]$. Let $m(i)$ be the population at location i. Assume that $m(\cdot)$ is normalized so that total population in the economy is $\int_0^1 m(i)di = 1$. There is also a continuum of commodities indexed by the pair (x, k), where $x \in [0,1]$ and $k \in \{1,2\}$. See Figure 1 for an illustration of the important characteristics of the set of commodities. The set of commodities is a pair of line segments. The k selects the line segment while the x indicates where we are at along the line segment. Specifically, the x indexes the type of good, e.g., grocery wholesaling or cabinet making, while the k indexes whether must be locally produced ($k = 1$) or can possibly be imported ($k = 2$).[3] We will refer to these as Type 1 and Type 2 goods. The inability to ship Type 1 commodities can be thought of as Type 1 goods possessing infinite transportation costs. Type 2 commodities can be transported at a cost that varies with the sector to which the good belongs. The goods are classified into three sectors: manufacturing, services and retail. Let μ, σ, ρ, be the fraction of industries in the three sectors, $\mu + \sigma + \rho = 1$. Assume the goods are ordered such that goods $x \in [0, \mu)$ are manufactured goods, $x \in [\mu, \mu + \sigma)$ are services, and $x \in [\mu + \sigma, 1]$ are retail goods. Make the extreme assumption that manufactured goods have zero transportation cost, $\tau_M = 0$, retail transportation costs are prohibitive, $\tau_R = \infty$, and service transportation costs are intermediate, $0 < \tau_S < \infty$. Transportation costs are modeled in the usual "iceberg" way. In order to deliver a type x good to another location, $1 + \tau_x$ units must be shipped; τ_x units get lost along the way.

2.2 Consumers

Assume that individuals have Cobb-Douglas preferences represented by the utility function,

$$U\left(\{q\left(x,k\right) : x \in [0,1], k \in \{1,2\}\}\right) = \int_0^1 \left[\lambda \ln q(x,1) + (1-\lambda) \ln q(x,2)\right] dx$$

where $q(x, k)$ denotes the consumption by an individual of commodity (x, k). In equilibrium, the preference parameter λ will denote the share of income that is spent on goods that must be locally produced. Consumers have one unit of labor which is inelastically supplied to

[3] Another way to think about this is that for $k = 1$ goods, consumers exhibit an extreme home bias in consumption, while for $k = 2$ goods they have absolutely no home bias.

firms. Consumers take prices and the wage rate as given, and choose consumption to maximize utility subject to a budget constraint,

$$\int_0^1 \left[p\left(x,1\right)q\left(x,1\right)+p\left(x,2\right)q\left(x,2\right)\right]dx \leq w$$

Strictly speaking, the prices, wage rate, and quantities consumed also depend upon the location, but we suppress the dependence for now. The consumer also receives a share of the firms' profits, but since profits will be zero in equilibrium they are not shown here.

2.3 Producers

Labor is the only factor of production. Firms can choose to use one of two technologies.[4] The first technology is constant returns to scale. For this technology, one unit of output requires γ units of labor as input. The second technology has a range of increasing returns to scale. For the scale economy technology, q units of output requires $c(q)$ units of labor input. Define the average cost (in labor units) to be $a(q) = c(q)/q$. Assume that average cost is minimized at a point $q^* \in (0, \infty)$ where q^* is defined as the lowest point the minimum is attained (i.e. q^* is the minimum efficient scale). The average cost function can be either U-shaped or can flatten out for $q > q^*$ so that average cost is constant beyond q^*. Assume that for $q < q^*$, marginal cost $c'(q)$ is strictly increasing. Define $a^* \equiv a(q^*)$. Figure 2 displays the cost curves associated with these technologies.

We assume that the scale technology has a lower average cost at the minimum efficient scale than the constant returns technology,

$$\gamma > a^*. \tag{A1}$$

This assumption avoids the trivial equilibrium in which all firms use the constant returns technology and each location produces its own requirements of all the goods. We also make an assumption restricting the population size of the locations we will consider. Specifically,

[4] This is not the same as in Krugman (1991) where there were two technologies corresponding to different goods: a scale technology for the differentiated goods and a constant returns technology for the homogeneous good. In our model each good has two possible production methods, but all goods have the same two possibilities.

we assume that

$$\ell < c\left(\ell/\gamma\right). \tag{A2}$$

This ensures that if a scale firm exists, it will be the only firm at that location producing that particular good.[5] It also implies that a firm using the scale technology must export to survive. Without this assumption, equilibria in which the scale firm coexists with other firms (using the constant returns or scale technology) may exist, complicating the analysis without, we believe, altering the basic results. Define

$$\ell_{\max} \equiv \min\left\{\ell : \ell = c\left(\ell/\gamma\right),\ \ell > 0\right\}$$

and

$$L \equiv \left(0, \ell_{\max}\right).$$

This is the set of population levels admissible under assumption (A2). Given our assumptions on the cost function, L is non-empty.

Only manufacturing and service goods can be exported as retail goods are assumed to have infinite transportation costs. Let the $(x, k) = (0, 2)$ good, a Type 2 manufactured good, be defined as the numeraire, and let $p(x, k, i)$ be the local price, including transportation costs, of commodity (x, k) at location i in terms of the numeraire good. All firms are price-takers in export markets. Let $p^E(x)$ be the export price for good x. The export price is quoted as the mill price; it does not include the transportation cost. Firms that use the constant-returns to scale technology are price takers in the local market, while firms that use the scale-economy technology are price setters in the local market. Assumption (A2) ensures that there is at most one scale-economy firm at each location for any particular good. With the price of Type 2 manufactures normalized to 1, this will be both the local and export price for manufacturing goods since there are no transportation costs for these goods. That is, if good x is a manufacturing good, then $p(x, 2, i) = p^E(x) = 1$ for every location i. The price of Type 1 manufacturing goods, $p(x, 1, i)$, will exceed 1 in equilibrium.

[5] Rearranging (A2) we get $\gamma < c\left(\ell/\gamma\right)/\left(\ell/\gamma\right)$. This says that to produce the quantity ℓ/γ (the demand that would prevail in autarky), the average labor requirement with the CRS technology is lower than that for the scale technology. Hence, the scale technology will not be used unless exports are possible.

Finally, all firms behave competitively in the labor market. Let $w(i)$ be the wage at location i.

A firm using the scale technology produces for three markets: the local Type 1 market, the local Type 2 market, and the export market. Denote the quantities produced for these markets by $q^1\left(w, p^E, \ell\right)$, $q^2(w, p^E, \ell)$, and $q^E(w, p^E, \ell)$. From the firm's perspective these quantities are perfect substitutes in production, so let

$$\tilde{q}(w, p^E, \ell) = q^1(w, p^E, \ell) + q^2(w, p^E, \ell) + q^E(w, p^E, \ell)$$

be the total quantity produced. To conserve notation, for a given location i with population ℓ and a given good (x, k), we define $p^1 \equiv p(x, 1, i)$, $p^2 \equiv p(x, 2, i)$, $p^E \equiv p^E(x)$, and $w \equiv w(i)$, where p^1 and p^2 denote the local prices for Type 1 and Type 2 good x, p^E is the export price taken as given by firms in each location, and w is the wage rate. Then, given p^E, w, and ℓ, firms maximize the profit function

$$H\left(w, p^E, \ell\right) = p^1 q^1 + p^2 q^2 + q^E p^E - c(\tilde{q}(w))w \quad (1)$$

by setting the two local prices, p^1 and p^2, and choosing a total quantity to produce, $\tilde{q}(w)$.

2.4 Equilibrium

An *equilibrium* for this model is a set of price functions $p(x, k, i)$, $p^E(x)$, $w(i)$, a set of consumption decisions, a set of entry decisions of firms whether to produce a particular good at a particular location using the scale-economy technology, and output decisions by firms such that the following conditions hold. First, consumers maximize utility given wage income and any share of firm profits (in equilibrium these profits turn out to be zero). Second, firms that enter maximize profit. Third, it is not profitable for another firm to enter given the entry and output decisions of other firms. Fourth, supply equals demand for each sector.

3 Characterization of Equilibrium

This objective of this section is to prove the existence and uniqueness of equilibrium in this model and to characterizes its properties. Our main result is summarized by Proposition 1 which claims that in addition to being unique, the equilibrium in this model is of a particular type, namely, a *zero-profit specialization equilibrium.*

Proposition 1 *There exists a unique equilibrium in this model. This equilibrium is characterized by firms earning zero-profits and locations specializing in a particular sector for its export goods.*

We prove this proposition by constructing an equilibrium which turns out to be the *unique* equilibrium of the model. Let us briefly consider how each element of the model is solved for.

- The immobile labor supply at each location is assumed to be exogenous.
- Start by taking the wages at each location and the export price as given. (Recall that the manufacturing price is one in all markets.)
- Solve the representative consumer's utility maximization problem to obtain the local market demand functions.
- Producers maximize profits by setting the local Type 1 and Type 2 prices to keep out constant returns to scale establishments and imports respectively. The export price determines the total quantity produced as the marginal unit is sold on the export market. Then, given the local quantities demanded, we can solve for the amount that the producer exports.
- Given the export price and free entry by firms, each location's sector of specialization is determined by a population cut-off rule. Except for locations with a population level equal to the cut-off (a set of measure zero if the distribution over population sizes is continuous), one sector will always be able to out-bid the other for labor.
- Given a sector of specialization for each location, the wage rate and the measure of firms using the scale technology at each location are determined jointly by the zero-profit condition, as there is free entry and exit, and the labor market clearing condition.[6]

[6]The measure of firms using the scale technology at a location is equivalent to the measure of varieties produced at that location using the scale technology. This is a consequence of assumption (A2) which ensures that there exists at most a single scale firm producing a given product at any location.

- Market clearing in the service sector pins down the export price $p^E(x)$. The manufacturing sector clears by Walras law.

- Finally, we know the measure of firms producing in each sector at each location, but we do not know which specific goods are produced at each location. An assignment algorithm can be used to map goods varieties and locations.

Notice that while a location's sector of specialization *is* unique, the specific varieties produced at that location are not uniquely determined. That is, except for the assignment algorithm, the equilibrium is unique. This sort of indeterminacy arises in other models with scale economies in the production of differentiated goods (eg., Krugman (1980)). As in those models, this indeterminacy does not have any effect on the analysis.

For now we will consider a representative good x produced at location i. Start by taking wages and export prices as given and solve for the equilibrium prices and quantities demanded in the local market. Bertrand competition ensures that a firm using the scale technology would set a price equal to that of a firm using the constant returns technology. Therefore, the price and total quantity demanded for Type 1 goods at a location with population ℓ are

$$p^1 = \gamma w$$

and

$$q^1 = \frac{\lambda w \ell}{p^1} = \frac{\lambda \ell}{\gamma}$$

The firm is a price setter in the local market for Type 2 goods, so it sets a price equal to the delivered price at other locations,

$$p^2 = p^E(1+\tau)$$

The quantity demanded at this price is

$$q^2 = \frac{(1-\lambda)\, w\ell}{p^2} = \frac{(1-\lambda)\, w\ell}{p^E(1+\tau)}$$

The price p^2 can also be thought of as the outcome of Bertrand competition between the local firm and potential exporters in other locations. Finally, in the export market firms are price takers, so profit maximization implies output must solve the following first-order necessary condition.

$$p^E = c'(q)w.$$

This condition implicitly defines a function, $\tilde{q}(w, p^E)$, that gives the optimal total output as a function of the export price and the local wage rate. This function is strictly decreasing in the wage rate and strictly increasing in the export price.

Figure 3 illustrates this case for a representative service sector establishment. The marginal and average cost curves are standard. The marginal revenue, however, takes on a stair-step shape since the establishment sells to three markets. The first horizontal segment corresponds to the marginal revenue from sales to the local Type 1 demand at the price γw. The second horizontal segment corresponds to the marginal revenue from sales to the local Type 2 demand at the price $p^E(1+\tau)$. The third segment is the marginal revenue from export market sales at the price p^E. Several points are in order regarding this figure. First, recall assumptions (A1) and (A2). The role played by (A1) is to ensure that scale technology is preferred to the constant returns technology in some neighborhood of the MES. If this were not the case, only the CRS technology would be used and one price would prevail for all goods at all locations. Assumption (A2) ensures that the marginal cost and marginal revenue curves cross in the third segment of marginal revenue. Second, notice that the chosen quantity, $\tilde{q}\left(w, p^E\right)$, falls to the left of the minimum efficient scale. That is to say, local market power leads productive inefficiency on the part of firms using the scale technology. Finally, if we set $\tau = 0$ and $p^E = 1$ as is the case for a manufacturing establishment, then the marginal revenue collapses to two line segments rather than three. Even so, productive efficiency is still not achieved as the manufacturing firm retains its price-setting power in the local Type 1 market.

Using the local demand functions, the profit function (1) can be rewritten as,

$$H(w, p^E, \ell) = w\ell + \left[\tilde{q}(w, p^E) - \frac{\lambda\ell}{\gamma} - \frac{(1-\lambda)\,\ell w}{(1+\tau)p^E}\right] p^E - c(\tilde{q}(w, p^E))w$$

In the event of positive profits, free entry by firms increases the demand for labor, driving wages up until profits are equal to zero. Let $w(p^E, \ell)$ denote the wage rate which solves the zero profit condition, $H(w, p^E, \ell) = 0$.

It proves useful to define two particular wage levels. First, define $w^*\left(p^E\right)$ to be the wage such that production occurs at precisely the minimum efficient scale, $\tilde{q}(w^*) = q^*$. At this point,

$$p^E/w^* = a^* = c(q^*)/q^* = c'(q^*). \tag{2}$$

Since w^* depends on p^E, it will therefore differ between locations specializing in services and those specializing in manufactures. Next, define $w'\left(p^E, \ell\right)$ by,

$$\tilde{q}(w', p^E) \equiv \frac{\lambda \ell}{\gamma} + \frac{(1-\lambda)\,\ell w'}{(1+\tau)p^E} \tag{3}$$

This is the wage level such that exports drop to zero. Under assumption (A2), $w'\left(p^E, \ell\right) > w^*\left(p^E\right)$ and therefore, $\tilde{q}(w^*, p^E) > \tilde{q}(w'(\ell), p^E)$ as $\tilde{q}(\cdot)$ is strictly decreasing in w. Our next claim serves to restrict the set of wages and prices that must be considered when solving for an equilibrium with all wages and prices determined endogenously.

Lemma 1 *In equilibrium, the wage consistent with zero-profits at each location and the export price for services* $\left(p^E\right)$ *must satisfy the following:*

1. *(Wage is above the MES wage)* $w\left(p^E, \ell\right) > w^*\left(p^E\right)$

2. *(Wage is below the no-exports wage)* $w\left(p^E, \ell\right) < w'\left(p^E, \ell\right)$

3. *(Services do not dominate manufactures at all locations)* $p^E < 1$

4. *(Constant returns technology does not dominate at all locations)* $\gamma w\left(p^E, \ell\right) > (1+\tau)\,p^E$ *for some location not exporting manufactures and* $\gamma w\left(p^E, \ell\right) > 1$ *for some location not exporting services.*

Proof of all lemmas can be found in the appendix. The first two conditions in Lemma 1 restrict the set of wages at each location; they follow directly from free entry and exit and assumption (A2). These constraints on the admissible set of wages help to establish the next result on the existence and uniqueness of a wage for which profits are identically zero.

Lemma 2 *For a given* p^E *satisfying Lemma 1 and* $\ell \in L$*, there exists a unique* $w \in \left(w^*\left(p^E\right), w'\left(p^E, \ell\right)\right)$ *such that* $H(w, \ell) = 0$.

Having established the existence of a unique zero-profit wage for any given population level, we now ask how the zero-profit wage varies with population. The finding, in Lemma 3, is that wages increase with population for the set of population levels consistent with the existence of a single exporting firm.

Lemma 3 *Given p^E, the wage consistent with free entry and exit, $w\left(p^E, \ell\right)$, is increasing for $\ell \in L$.*

This is due to the fact that larger locations are relatively more attractive to firms with more price setting power. In our model, price setting power is greater among service sector establishments than among manufacturing establishments as their high transportation costs offer more protection from imports than is the case for manufacturing establishments. By price discriminating over a larger local base, these service sector firms have higher profits which attracts more firms and drives up wages. This is the classic home market effect (Krugman (1980, 1991)).

Let $w_S\left(p^E, \ell\right)$ denote the zero-profit wage for services and let $w_M(\ell)$ denote the zero-profit wage for manufacturing (recall that the manufacturing export price is one). These functions are analogous to bid-rent functions in urban economics (see Mills and Hamilton (1994)). Figure 4 plots these functions. The activity that supports the highest wage is the activity that the location specializes in.

For manufacturing, access to low-cost labor is more important than proximity to markets, while for services, market access is relatively more important. If the service sector export price, p^E, were equal to the manufacturing export price of one, then the graphs of $w_S\left(p^E, \ell\right)$ and $w_M(\ell)$ would start at the same point. Because services have higher transportation costs, they would always be more profitable than manufacturing in the local markets, and equally profitable in export markets. Therefore, due to greater profitability, the service establishments would bid wages up higher than manufacturing and graph of the zero-profit wage for services would remain higher than that for manufacturing. In this case, no manufacturing goods would be produced using the scale technology, a situation that cannot persist in equilibrium as discussed in Lemma 1. For lower service sector export prices, the graph of the zero-profit wage for services starts below that of manufacturing and crosses the manufacturing graph at a positive level of labor, $\hat{\ell}$.

Lemma 4 *There exists a unique* $\hat{\ell} \in L$ *such that* $w_M\left(\hat{\ell}\right) = w_S\left(p^E, \hat{\ell}\right)$. *Locations with* $\ell < \hat{\ell}$ *specialize in manufacturing and locations with* $\ell > \hat{\ell}$ *specialize in services.*

Every location must produce Type 1 manufactures and Type 1 services (as well as both Type 1 and Type 2 retail goods). The production of non-retail Type 2 goods will take place in firms using the scale technology, and each location will have a single firm using the scale technology for any given product. Since locations differ in population, each location will have a measure of firms using the scale technology for distinct goods within the same sector. The determination of which locations produce Type 2 services and which produce Type 2 manufactures is made on the basis of wages as the immobile workers will work for that firm which offers the highest wage. The tension between market access and low-cost labor make lower transportation costs sectors better suited for small locations and high transportation cost sectors better suited for large locations.

Given export prices and the sector of specialization, firms enter and exit freely, bidding up wages until profits are driven to zero and the labor market clears. Labor supply at location i is simply $m(i)$. Let $n_z(i)$ denote the measure of scale technology firms at location i where $z \in \{M, S\}$ denotes the sector of specialization. In other words, n_z is the measure of the set of x's produced for export. Labor demand at location i has three components. First, for every non-exported good $x \in [0, 1]$, firms require $q(x, 1, m(i))\gamma$ workers to satisfy local demand for the Type 1 variety of the good. Second, for every Type 2 retail good, $x \in [\mu + \sigma, 1]$, firms require $q(x, 2, m(i))\gamma$ workers to satisfy the local demand for that good. Finally, for the set of x, $x \in [0, \mu)$ or $x \in [\mu, \mu + \sigma)$ depending upon the sector of specialization, $n_z(i)\, c(q(x))$ units of labor are required to meet the Type 1 and Type 2 demand (including export demand) for that good. Let $I_{(i,x)}$ be an indicator function that equals one if location i produces good x using the scale technology and equals zero otherwise. Then, labor clearing in location i is

$$\begin{aligned} m(i) &= \int_0^{\mu+\sigma} \left[q(x, 1, m(i))\gamma\left(1 - I_{(i,x)}\right) + c(\tilde{q}(x))\, I_{(i,x)}\right] dx \\ &\quad + \int_{\mu+\sigma}^{1} \left[q(x, 1, m(i))\gamma + q(x, 2, m(i))\gamma\right] dx \end{aligned}$$

$$= (\mu + \sigma - n_z(i))\,\lambda m(i) + n_z(i)\,c(\tilde{q}(x^*)) + (1 - \mu - \sigma) m(i)$$
$$= m(i)\left[(\mu + \sigma - n_z(i))\,\lambda + (1 - \mu - \sigma)\right] + n_z(i)\,c(\tilde{q}(x^*))$$

where the first integral corresponds to manufacturing and service labor demand and the second integral is the retail labor demand. Since each location specializes in a single sector and the price for all goods within a sector are the same, it is without loss of generality that we let x^* represent a typical good produced for export in location i. The second equality then uses the quantity of the representative x produced and the measure of x's produced using the scale technology to calculate the required labor. Solving for the measure of firms we get

$$n_z(i) = \frac{(1 - \lambda)(\mu + \sigma) m(i)}{c(\tilde{q}(x^*)) - \lambda m(i)}.$$

This expression says that the measure of scale sector firms is equal to the total labor "allocated" to Type 2 manufactures and services production divided by the total scale firm demand for labor less that labor which, in the absence of a scale sector firm for this good, would be used exclusively for Type 1 production. Recall that $\tilde{q}$ depends on wages, so lower wages will increase $\tilde{q}$ thereby lowering the measure of scale technology firms at a given location.

The characterization of equilibrium so far has assumed that the export price is constant. We now consider the full equilibrium, taking into account trade between regions and the determination of the service sector export price, p^E. The graph of $w_S\left(p^E, \ell\right)$ shifts down and to the right with decreases in the export price as lower export prices imply less profit and therefore less pressure on wages (see figure 4). Equivalently, lower export prices imply that the firm needs a larger local base on which to price-discriminate. Thus, varying the export price will vary the population cut-off. For example, as we lower p^E, more locations will specialize in manufacturing which decreases the supply of services relative to manufactures. Thus, by consider the net supply of services to the export market, we can determine the price needed for market clearing in the service sector. That is, to solve for the equilibrium service sector export price we need the sum of net exports across locations to be exactly zero. Market clearing in the manufacturing sector will follow by Walras' law. Let q^1 and q^2 refer to the Type 1 and local Type 2 demands for a representative service good. Then the demand for *all* services at location i is simply the measure of service goods, σ, times the local demand for a representative good,

$$Q_S^D\left(i,p^E\right)=\sigma\left[q^1\left(w,p^E,m\left(i\right)\right)+q^2\left(w,p^E,m\left(i\right)\right)\right].$$

Location i's total supply of services is

$$\begin{aligned}Q_S^S\left(i,p^E\right) &= n_s\left(i\right)\left[q^1\left(w,p^E,m\left(i\right)\right)+q^2\left(w,p^E,m\left(i\right)\right)+q^E\left(w,p^E,m\left(i\right)\right)\right]\\ &\quad +\left(\sigma-n_s\left(i\right)\right)q^1\left(w,p^E,m\left(i\right)\right)\\ &= n_s\left(i\right)\left[q^2\left(w,p^E,m\left(i\right)\right)+q^E\left(w,p^E,m\left(i\right)\right)\right]+\sigma q^1\left(w,p^E,m\left(i\right)\right)\end{aligned}$$

where $n_s\left(i\right)$ is the measure of service sector firms using the scale technology. The net supply of services (i.e., net exports) by location i is

$$\begin{aligned}Q_S^S\left(i,p^E\right)-Q_S^D\left(i,p^E\right) &= n_s\left(i\right)\left[q^2\left(w,p^E,m\left(i\right)\right)+q^E\left(w,p^E,m\left(i\right)\right)\right]-\sigma q^2\left(w,p^E,m\left(i\right)\right)\\ &= \left(n_s\left(i\right)-\sigma\right)q^2\left(w,p^E,m\left(i\right)\right)+n_s\left(i\right)q^E\left(w,p^E,m\left(i\right)\right).\end{aligned}$$

Market clearing requires

$$\int_0^1\left[Q_S^S\left(i,p^E\right)-Q_S^D\left(i,p^E\right)\right]di=0.$$

This can then be solved for p^E.

As discussed earlier, we know which locations produce for which export sector, but we do not know which locations produce which specific goods using the scale technology. While this indeterminacy has no bearing on the model, we offer the following algorithm as *one* possible means of assigning production of the goods to locations. Take the set of locations specializing in manufactures. Let the "first" location, $i=0$, produce goods $[0,n_m\left(i\right)]$. Let location $i+\varepsilon$ (recall that there are a continuum of locations) produce goods $[0,n_m\left(i+\varepsilon\right)]$. Continue in this fashion until enough locations are producing good 0. Suppose this happens after location i' produces goods $[0,n_m\left(i'\right)]$. Then location $i'+\varepsilon$ will produce goods $(0,n_m\left(i'+\varepsilon\right)]$. Note that good 0 is excluded now. As we move over the location index, the interval of goods produced shifts. Figure 5 provides an illustration of this correspondence. The vertical axis indexes the goods by sector and the horizontal axis indexes the locations. Locations to the left of the population cut-off $\hat{\ell}$ produce manufactures and those to the right produce services.

None of the locations produce retail goods (the highest transportation cost goods) with the scale technology.

The proof of Proposition 1 follows directly from Lemmas 1 through 4. Specifically, we have shown that there exists a unique equilibrium in which all firms earn zero-profits and each location specializes in a particular sector. The next section discusses the implications of this model economy.

4 Implications of the Model Economy

The following propositions formally state the model's implications and make it clear that this model is able to capture the empirical facts observed in the data.

Proposition 2 (Specialization) *The pattern of specialization over goods that vary only in transportation costs is determined by the size of the populations at the geographic locations.*

Proof. This is immediate from Lemma 4. ∎

Lower transportation cost goods have a lower threshold size at which production can profitably be commenced using the more productive scale technology. This is due to the fact that relatively more goods arrive at their destination (due to lower iceberg transportation costs). In other words, market access is easier for low transportation cost goods and so the location of production matters relatively less. Therefore, low transportation cost industries locate in low-wage locations, having been bid out of the larger, high-wage locations by the high transportation cost industries for which proximity to final demand is more crucial. This process of bidding for labor leads to a cut-off level of population such that locations below the cut-off specialize in low transportation cost goods and those above the cut-off specialize in high transportation cost goods. Notice that Proposition 2, together with our assumptions on transportation costs, $\tau_M = 0$, $\tau_R = \infty$, and $0 < \tau_S < \infty$, gives rise to the same specialization patterns we observe in the data. Specifically, large locations specialize in the high transportation cost service-type industries, small locations specialize in the low transportation cost manufacturing industries, and all location produce the highest transportation cost goods—retail.

Proposition 3 (Establishment Size) *Within an industry, establishments are larger at those locations specializing in that industry.*

Proof. See appendix. ∎

Thus, regardless of the industry we are considering, locations with a firm using the scale technology will have larger establishments than locations without a scale establishment as those with scale establishments will be exporting. With our transportation cost assumptions, this implies that service-type industries have larger establishments in big cities while manufacturing industries have larger establishments in small cities and rural areas. Again, this result is borne out in the data. Note that smaller locations have a smaller measure of establishments using the scale technology rather than smaller establishments. Therefore, establishments are still larger in the small location's sector of specialization. The only way to contradict this result is for a location to be so large that the scale technology can profitably be used to meet local demand without exporting. This possibility is ruled out by assumption A2.

Proposition 4 (Productivity) *Locations that specialize, and hence export, in a given industry exhibit a higher labor productivity (output per worker) than locations not specializing in that industry.*

Proof. This follows directly from the two technologies; output per worker is higher when using the scale technology. ∎

Since specialization corresponds to exporting in that industry, Propositions 3 and 4 imply that exporting establishments are larger and more productive. This result has empirical support in Bernard and Jensen (1995, 1999), as discussed in the introduction.

5 Conclusion

This paper is an initial attempt to begin integrating the study of regional level facts (here specialization) and plant level facts (here plant size). Our model can account for why large cities specialize in services rather than manufacturing and why the size of establishments in big cities is above the national average in services and below the national average in manufacturing. A crucial ingredient of our theory is that within each sector there are always some aspects of the industry (e.g. custom work) that need a local presence and small plants exist to provide this custom work. While this is a simple and straightforward explanation, the formal analysis has to deal with complexities, such as the fact that a firm can simultaneously sell in three markets: (1) local demand that could be substituted by imports (i.e., work that is not custom work), (2) local demand that cannot be substituted by imports, and (3) export demand.

If our explanation is on the right track, it suggests that measures of industry concentration (e.g. Ellison and Glaeser (1997)) may in some sense understate the degree of industry concentration. If we were to focus only on that portion of demand that can be met by traded goods, industries would be more highly concentrated than if we were to also include the small plants doing only custom, non-tradeable work. In Holmes and Stevens (forthcoming), we find with U.S. data that when we restrict attention only to large plants, industry concentration is significantly higher than when we include all plants.

Our model is a substantial departure from the usual Dixit-Stiglitz economic geography model. We believe our structure—which puts firms in the position of acting like a monopolist in the local market while behaving competitively in export markets—may prove useful in other contexts. It has the benefits of the tractability of competitive models, but still captures the insights of the new economic geography literature. In Holmes and Stevens (2002) we have already used a variant of this structure to reexamine the issues about the home-market effect raised by Davis (1998). Given the huge amount of interest that the economic geography literature has generated, and given the well-known limitations of Dixit-Stiglitz models (in terms of their flexibility and generality), we think there is some value to at least entertain or explore an alternative formulation, such as we do in this paper.

6 Appendix A: Proofs

Proof of Lemma 1. Suppose $w \leq w^*$, then for $\ell > 0$ firms are guaranteed positive profits, as Bertrand competition in the local Type 1 market (with potential constant returns to scale establishments) ensures profits since $\gamma > a^*$. This is inconsistent with free entry.

Suppose $w \geq w'$, then there are no exports as $\tilde{q}(\cdot)$ is strictly decreasing in w, and w' is defined as that wage level at which $\tilde{q}(\cdot)$ exactly equals local demand. This means that assumption (A2) is violated.

Suppose to the contrary that $p^E \geq 1$, then no firm would produce manufactures as service goods can be exported at a price at least as great as that of manufactures, and local Type 2 service goods sell for a strictly higher price than local Type 2 manufactures since transportation costs are positive for services. However, if $p^E \geq 1$, then there will be excess demand for manufactures and so this cannot be an equilibrium export price for services.

Suppose $\gamma w < 1$ for all locations. Then no manufacturing firm at any location will ever use the scale technology as the constant returns to scale firms can underprice them. Similarly for services if $\gamma w < (1+\tau)\, p^E$ at all locations. ■

Proof of Lemma 2. We first show that $H(w,\ell)$ is strictly decreasing in w for $w < w'(\ell)$. Taking the partial derivative of the profit function with respect to w,

$$\begin{aligned}
\frac{\partial H(w,\ell)}{\partial w} &= \ell + p^E\left[\tilde{q}' - \frac{(1-\lambda)\,\ell}{(1+\tau)\,p^E}\right] - \left[c'\tilde{q}'w + c(\tilde{q})\right] \\
&= \left[\frac{\tau+\lambda}{1+\tau}\right]\ell - c(\tilde{q}) \\
&< \ell - c(\ell/\gamma) \\
&< 0
\end{aligned}$$

where second equality follows from the firm's first order condition in the export market. The first inequality follows from the fact that $\left[\frac{\tau+\lambda}{1+\tau}\right] < 1$ and that ℓ/γ is less than $\tilde{q}$. [ℓ/γ is the local demand if the producer sold only to local demand and charged a price $p = w\gamma$. $\tilde{q}$ is evaluated at a lower wage rate and includes exports. Since $c(\cdot)$ is an increasing function $c(\tilde{q}) > c(\ell/\gamma)$.] The second inequality follows from assumption (A2). Thus $H(w,\ell)$ is strictly decreasing for $w < w'(\ell)$.

Next we show that $H(w^*,\ell) > 0$. We have that

$$\begin{aligned} H(w^*,\ell) &= w^*\ell + \left[q^* - \frac{\lambda\ell}{\gamma} - \frac{(1-\lambda)\,\ell w^*}{(1+\tau)p^E}\right] p^E - c(q^*)w^* \\ &= w^*\ell\left[1 - \frac{\lambda p^E}{\gamma w^*} - \frac{(1-\lambda)}{(1+\tau)}\right] + \left[p^E q^* - c(q^*)w^*\right] \\ &= w^*\ell\left[1 - \frac{\lambda}{(\gamma/a^*)} - \frac{(1-\lambda)}{(1+\tau)}\right] \\ &> 0 \end{aligned}$$

The last equality and the inequality follow from $\gamma w^*/p^E = \gamma/a^* > 1$.

Finally, we show that $H(w'(\ell),\ell) < 0$. The profit function is

$$\begin{aligned} H(w',\ell) &= w'\ell + \left[\tilde{q}(w') - \frac{\lambda\ell}{\gamma} - \frac{(1-\lambda)\,\ell w'}{(1+\tau)p^E}\right] p^E - c(\tilde{q}(w'))w' \\ &= w'\ell - c(\tilde{q}(w'))w' \\ &\leq w'\left[\ell - c(\ell/\gamma)\right] \\ &< 0 \end{aligned}$$

where the second equality follows from the definition of $\tilde{q}(w')$. The inequalities follow from the fact that $\ell/\gamma \leq \tilde{q}(w')$ and assumption (A2).

The conditions of the intermediate value theorem are satisfied, hence there exists of a unique $w \in (w^*, w'(\ell))$ such that $H(w,\ell) = 0$. ■

Proof of Lemma 3. Implicit differentiation of the profit function, $H(w,\ell)$, implies

$$\frac{dw}{d\ell} = \frac{p^E\left(\frac{\lambda}{\gamma}\right) - w\left(\frac{\tau+\lambda}{1+\tau}\right)}{\ell\left(\frac{\tau+\lambda}{1+\tau}\right) - c\left(\tilde{q}(w)\right)}$$

The denominator is clearly negative (see proof of Lemma 2). The numerator can be rewritten as

$$\frac{\lambda}{\gamma}\left[p^E - w\gamma\left(\frac{\tau+\lambda}{\lambda+\lambda\tau}\right)\right] < \frac{\lambda}{\gamma}\left[p^E - w\gamma\right] < 0$$

The second inequality follows from Lemma 1 and assumption (A1) which demonstrate that $w > w^* = p^E/a^* > p^E/\gamma$. Hence, $\frac{dw}{d\ell}$ is unambiguously positive for $\ell \in L$. ■

Proof of Lemma 4. This is shown in three steps (Figure 4 helps to visualize the proof). First, show that given ℓ, the slope of $w(\cdot)$ is greater for the higher τ good.

$$\frac{d}{d\tau}\left(\frac{dw}{d\ell}\right) = \frac{d}{d\tau}\left(\frac{p^E\left(\frac{\lambda}{\gamma}\right) - w\left(\frac{\tau+\lambda}{1+\tau}\right)}{\ell\left(\frac{\tau+\lambda}{1+\tau}\right) - c\left(\tilde{q}(w)\right)}\right) > 0$$

for $\ell \in L$, where the equality follows from Lemma 3. The inequality follows from the envelope theorem.

Second, we must show that $w_S(\ell)$ lies below $w_M(\ell)$ for ℓ small enough. To do so, we first show that as $w \longmapsto w^*$ it must be that $\ell \longmapsto 0$ in order to maintain zero-profits. As $w \longmapsto w^*$ we have that $q \longmapsto q^*$ and so profits are

$$\begin{aligned} H(w^*, \ell) &= \lambda \ell w^* \left[1 - \frac{c'(q^*)}{\gamma}\right] \\ &= \lambda \ell w^* \left[1 - \frac{a^*}{\gamma}\right] \end{aligned}$$

By assumption $a^* < \gamma$ so it must be that $\ell \longmapsto 0$ for profits to be zero. Recall that the profit maximizing conditions for manufacturing and services evaluated at w^* are

$$1 = c'(q^*)\, w_M^*$$

and

$$p_S^E = c'(q^*)\, w_S^*$$

respectively. Since $p^E < 1$ this implies that $w_S^* < w_M^*$. Therefore as ℓ goes to zero, the wage function for services lies below the wage function for manufactures.

Third, we must show that $w_S(\ell)$ lies above $w_M(\ell)$ for ℓ large enough. Suppose to the contrary that for all $\ell \in L$, $w_S(\ell) < w_M(\ell)$. In this case no location exports services. For this to be the case is must be that p^E is too low for service firms to profitably export, therefore the net supply of services is

$$Q_S^S\left(i, p^E\right) = 0$$

for all locations i. At the price p^E, however,

$$Q_S^D\left(i, p^E\right) > 0$$

for all locations i. Hence, this is not an equilibrium export price.

By the intermediate value theorem there exists a unique $\hat{\ell} \in L$ such that $w_M\left(\hat{\ell}\right) = w_S\left(\hat{\ell}\right)$.

■

Proof of Proposition 3. Let $\ell_1 > \hat{\ell} > \ell_2$. By Lemma 4 we know that locations 1 and 2 specialize in services and manufactures respectively. From Lemma 3 we know that

$w_1 > w_2$. Recall that $\tilde{q}_1(w_1) > \ell_1/\gamma$ as ℓ_1/γ is the local quantity demanded if the firm faced no threat of imports.

We first show that location 1's service labor requirements for a particular good, $c(\tilde{q}_1)$, are greater than location 2's service labor requirements for that good, $\lambda\ell_2$, even if we assume location 2 has a single constant returns to scale service firm for that good. Thus,

$$c\left(\tilde{q}_1\left(p^E, w_1\right)\right) > c\left(\frac{\ell_1}{\gamma}\right) > \ell_1 > \lambda\ell_2$$

where the second inequality follows from assumption (A2).

We next show that location 1's manufacturing labor requirement for a particular good, $\lambda\ell_1$, is less than location 2's manufacturing labor requirements for that good, $c(\tilde{q}_2)$, even if we assume location 1 has a single constant returns to scale manufacturing firm for that good. Thus,

$$c(\tilde{q}_2(w_2)) > c(\tilde{q}_1(w_1)) > c\left(\frac{\ell_1}{\gamma}\right) > \lambda\ell_1$$

where the first inequality follows from the fact that location 1 has a lower wage. ■

7 References

Amiti, M. (1998) Inter-Industry Trade in Manufactures: Does Country Size Matter? *Journal of International Economics,* 44: 231-255.

Bernard, A.B., Jensen, J.B. (1999) Exceptional Exporter Performance: Cause, Effect, or Both? *Journal of International Economics,* 47: 1-25.

Bernard, A.B., Jensen, J.B. (1995) Exporters, Jobs, and Wages in U.S. Manufacturing, 1976-1987. *Brookings Papers on Economic Activity, Microeconomics,* 67-112

Davis, D.R. (1998) The Home Market, Trade, and Industrial Structure. *American Economic Review*, 88: 1264-1276.

Dixit, A.K., Stiglitz, J.E. (1997) Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, 67: 297-308.

Ellison, G., Glaeser, E. (1997) Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy, 105: 889-927.*

Fujita, M., Krugman, P., Venables, A.J. (1999) *The Spatial Economy: Cities, Regions and International Trade.* Cambridge, MA: MIT Press.

Holmes, T.J. (1999) Scale of Local Production and City Size. *American Economic Review Papers and Proceedings.,* 89: 317-20.

Holmes, T.J., Stevens, J.J. (forthcoming) Geographic Concentration and Establishment Scale. *Review of Economics and Statistics.*

Holmes, T.J., Stevens, J.J. (2002) The Home Market and the Pattern of Trade: Round Three. Manuscript, University of Minnesota.

Krugman, P. (1991) Increasing Returns and Economic Geography. *Journal of Political Economy* 99: 483-499.

Krugman, P. (1980) Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review* 70: 950-959.

Melitz, M.J. (1999) The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Manuscript, University of Michigan.

Mills, E.S., Hamilton, B.W. (1994) *Urban Economics*, Fifth Edition. New York: Harper-Collins.

U.S. Department of Commerce (2001) Economic Census 1997, CD-rom Disk C1-E97-NA1F-17-US1.

Figure 1: Commodities

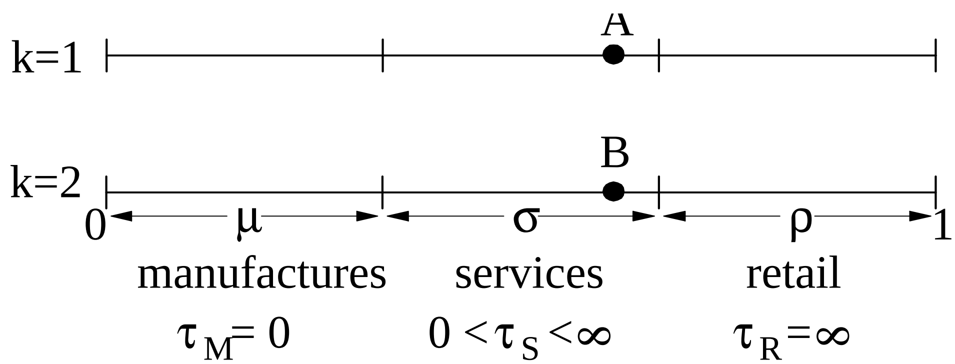

Figure 2: Available Technologies

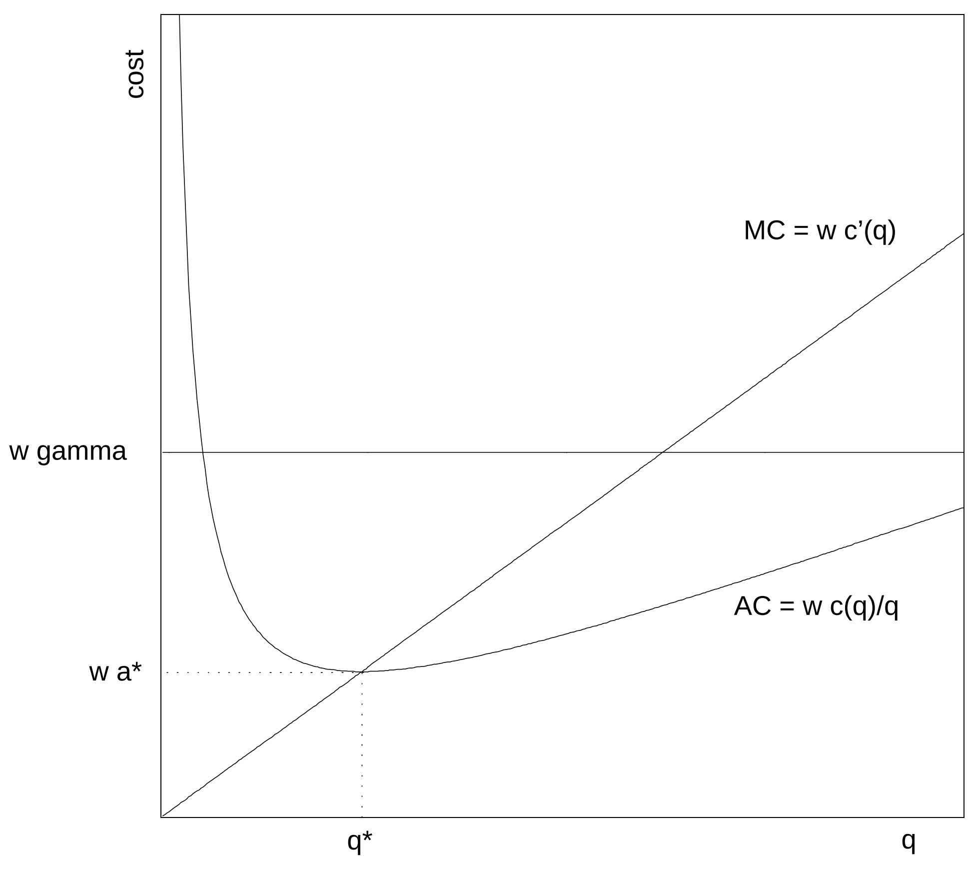

Figure 3: A Representative Service Sector Establishment

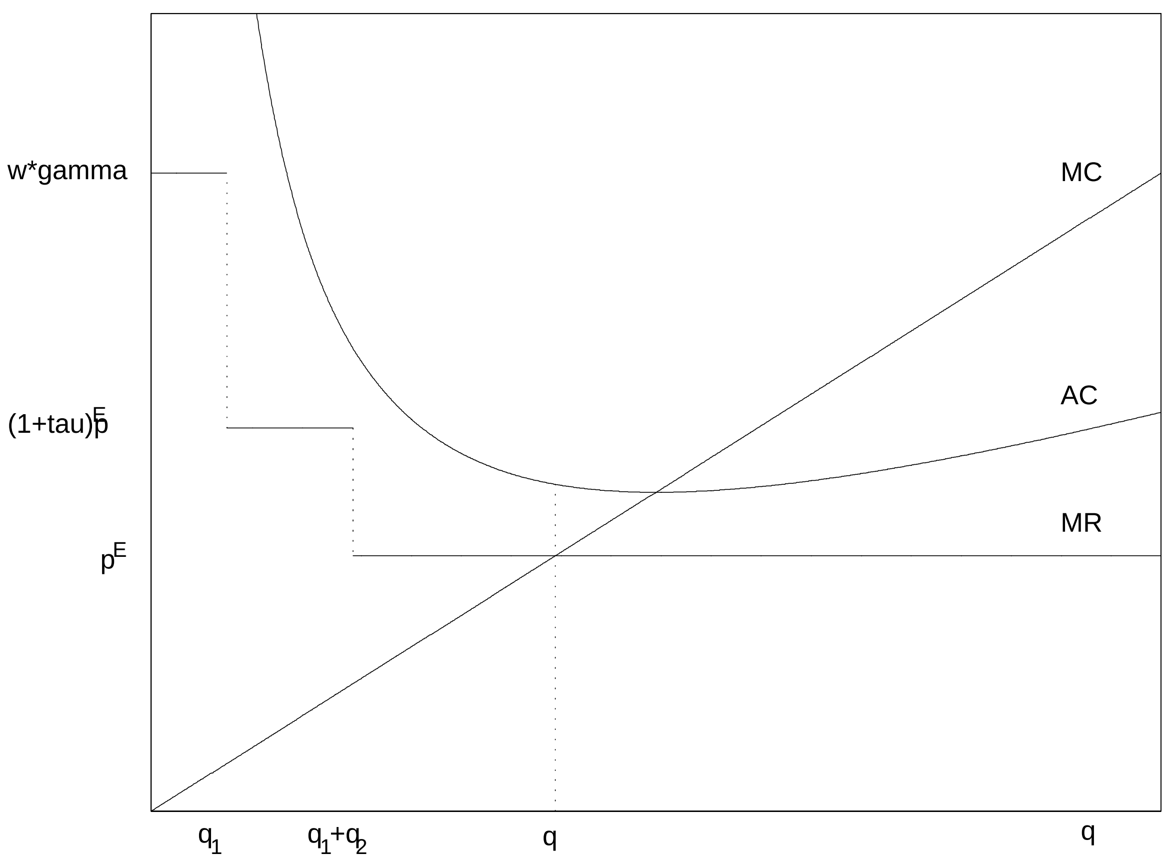

Figure 4: Bid-Wage Function and Specialization

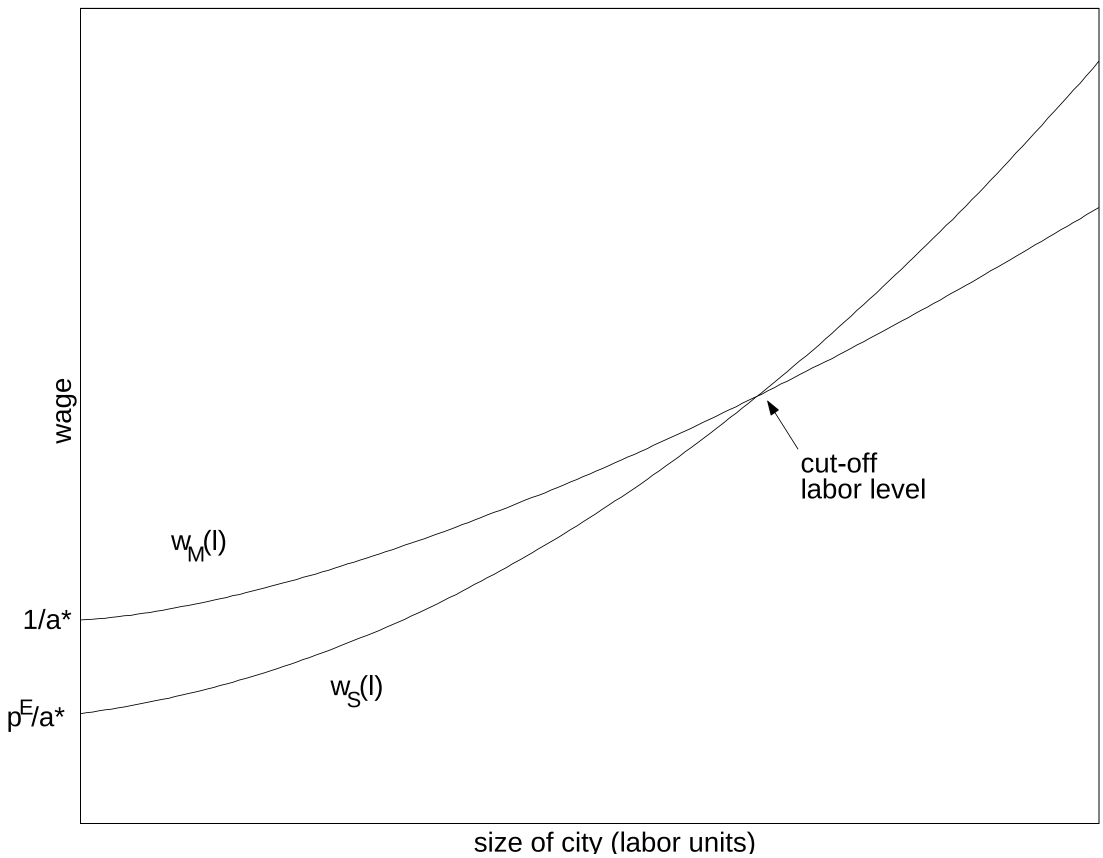

Figure 5: Correspondence Assigning Goods To Locations

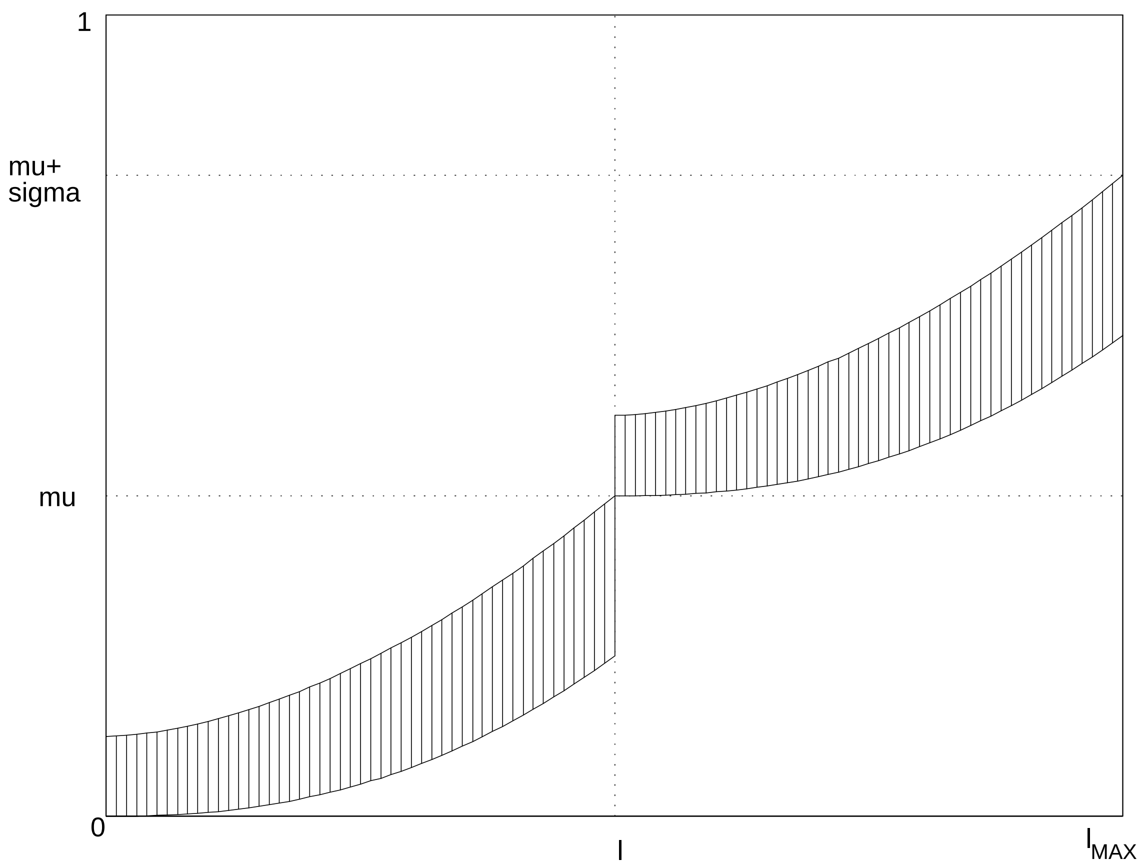

Table 1
City Specialization and Establishment Size
By City Size and Major Industry

Industry	Location Quotient			Size Quotient		
	Small Cities	Medium Cities	Large Cities	Small Cities	Medium Cities	Large Cities
Service-type Sectors						
Wholesale Trade	.54	.97	1.27	.70	.98	1.12
Finance and Insurance	.52	.89	1.36	.54	.86	1.38
Professional, Scientific, and Technical Services	.49	.87	1.40	.64	.87	1.23
Manufacturing	1.15	1.06	.86	1.25	1.08	.81
Retail	1.01	1.04	.96	.91	1.03	1.04

Source: Authors calculations with geographic data from the 1997 Economic Census (U.S. Department of Commerce (2001)). City size is defined by 1997 PMSA population: small (under 500,000), medium (500,000 to 2 million), and large (over 2 million). The Location Quotient is the share of sales in the city size class divided by the share of population. The Size Quotient is that average establishment sales size at the location divided by average U.S. sales size.