FORECASTING THE LOCATION OF INDUSTRIES*

Curtis C. Harris, Jr.

University of Maryland

Recently I presented a model for forecasting economic activity for all of the approximately 3,000 county-type areas in the United States.[1] The principal component of the overall model is a set of equations explaining the location of industries. It was pointed out that improvements could be made in these equations, but the improved equations were not derived. The purpose of this paper is to present the derivation of this new set of industrial location equations.

Traditionally, industrial location theory has been concerned with the optimum location of a plant (usually with respect to transportation costs) on a point in a continuous plane. More realistically, firms first choose from a finite set of broadly defined regions, such as states, counties, and metropolitan areas. Once the region has been selected, sites within the region are considered. Again though, only a finite number of them are considered. We are concerned with the maximum profit locations within a set of finite regions (counties). Moreover, we are dealing with the location of broadly defined industries and not individual plants. Our location equations, however, are consistent with the location of a firm.

In the first part of this paper the general form of the equations is derived from production theory. Next, a transition is made from the general equations to the equations to be estimated. Then, some of the explanatory variables are examined in more detail.

The General Form of the Locations Equations

Assume that the implicit production function of a firm is given as

$$F(q_1, \ldots, q_n) = 0, \tag{1}$$

where q_i denotes the quantity of the i^{th} output or input. Inputs are distinguished by a negative sign.

If the firm were in a pure competition situation, prices of these quantities of output and input would be given. However, as economists interested in location and space problems have pointed out, prices of some outputs and inputs are not given, but vary, both by the firm's location vis-a-vis its markets and by the volume of shipments. In order for a firm to sell additional units of output, it must incur additional transportation costs of having its products shipped greater distances. Higher prices also are associated with obtaining additional material inputs. Either the transportation costs or the factory prices go up as additional units are demanded. Prices are a function of distance, but the distance that goods are shipped is related to output. Therefore, prices are a function of output.

Rather than assuming prices given, we assume that prices are a function of output, as

$$P_i = p_i(q_i), \tag{2}$$

where P_i denotes prices of the i^{th} output or input. Revenue or cost of the i^{th} output or input is

$$C_i = p_i(q_i)q_i. \tag{3}$$

Solving (3) for q_i we get $q_i = Q_i(C_i)$. Substituting into (1), we have

$$F[Q_i(C_i), \ldots, Q_n(C_n)] = 0. \tag{4}$$

Let us denote the function of the C's on the left of (4) by

$$G(C_1, \ldots, C_n) = F[Q_1(C_1), \ldots, Q_n(C_n)] = 0. \tag{5}$$

Next, we want to see how the change in cost of one of the inputs or outputs is related to the changes of the others. By total differentiation of (5) we have

(6) $$G_1 dC_1 + G_2 dC_2, \ldots, G_n dC_n = 0,$$

where G_i is the partial differential of (5) with respect to C_i.

Now assume that the firm is a one-product firm where the n^{th} C represents revenue and the remaining (n-1 = m) C's represent costs of the various types of inputs. Let:

$$R = C_n$$

and

$$\beta_i = G_i/G_n. \quad (i = 1, \ldots, m).$$

Now (6) can be rewritten as

(7) $$dR = \sum_{i=1}^{m} \beta_i dC_i.$$

Although (7) was derived from a firm's production function, it is the general form of the industry location equations. We now derive the estimating equations.

<u>The Estimating Equations</u>

The changes in costs as given in (7) are assumed to be costs of acquiring marginal units of the inputs. Let marginal cost (MC) be defined as

(8) $$MC_i = dC_i/dq_i \quad (i = 1, \ldots, m).$$

Solving (8) for dC_i, substituting into (7), and adding a regional subscript, j, we have

(9) $$dR_j = \sum_{i=1}^{m} \beta_i MC_{ij} dq_{ij} \quad (j = 1, \ldots, r),$$

where r is the number of regions. Equation (9) assumes that the production functions and price functions are the same for each region.

In order to estimate the β coefficients in (9), we assume that dq_{ij} is equal to one;[2] and we use least-squares estimating procedures with regional data as observations. The regional variations in the marginal costs are being used to explain the regional variation in the change in value of shipments. In other words, we assume that the change in cost of an input as given in (7) is the cost of acquiring a marginal unit of the input.

Since we will have annual data, the changes in revenues and costs will be annual changes. Many, if not all, regions will have some changes during the year. (This refers to all changes, not just those associated with changes in capacity.) Thus, we are not finding the optimum location of an additional unit, but explaining the location of many additional units. However, we can still think of the locational decisions as being made for one unit (or a small number of units) at a time. The changes noted during the year are the cumulative results of such decisions. At the beginning of a period there is a set of marginal costs for each region. Then, given the location of the marginal unit of demand, the production of this unit will take place in the location where marginal profits are maximized. Once this production decision has been made, then the values of the marginal costs associated with each region will change. Thus, the maximum profit location of the next units might be somewhere else. This may be true even when the additional demand is located in the same place. It certainly is true when the additional demand is at another location. In the latter situation, the maximum profit producing location will be different, even if the marginal costs do not change. When we observe the changes during the year, production changes in many different places, reflecting not only the demand at the various locations but also changes in the marginal cost.

The marginal cost from several previous time periods may inluence the changes in production observed during the current time period. Some of these changes in production are a result of changes in capacity, and decisions to locate capacity are part of the overall investment decision. The planning-construction period associated with added capacity may involve several years. Thus, production coming from new investment in the current period may be a result of a decision made several years ago, and this decision would have been made in accordance with the marginal costs that existed at that time. In essence, what we have is a distributed lag situation in which investment decisions are made on expected costs and these expected costs are a function of existing costs. The equations that will be used for forecasting may include marginal costs from several previous time periods, either entered as separate variables or combined giving appropriate weights to each year.

It is recognized that the relationship between change in the value of shipments and the marginal costs may not be linear. If the non-linear relationship is hypothesized, then non-linear forms will be used, such as logarithmic forms, quadratic functions, or dummy variables.

The Explanatory Variables

For the convenience of discussion, let us divide the types of cost into four groups: (1) cost of materials, (2) transportation costs, (3) labor costs, and (4) other costs. The cost of materials is expressed in producers prices at the point of shipment and does not include transportation cost. Output expressed in physical quantity terms would be impossible, since we are working with broad industry classifications.

Cost of Materials: Since the cost of materials is expressed in producers prices, the marginal cost of an additional dollar's worth of material is, of course, equal to one. Equation (8) would be

$$MC_i = dC_i/dC_i \quad (i = 1, \ldots, s) \tag{8a}$$

where s is the number of material inputs. When dC_i is assumed to be one, the sum of the β coefficients of these costs is a constant. Let:

$$a = \sum_{i=1}^{s} \beta_i MC_{ij} dC_{ij} = \sum_{i=1}^{s} \beta_i \quad (j = 1, \ldots, r)$$

Transportation Costs: The costs having the most regional variation are transportation costs of shipping both the outputs and the inputs. The procedure used for estimating the marginal transportation costs is to compute shadow prices from a linear programming transportation problem.[3] There are shadow prices both for getting a marginal dollar's worth of product to a particular region, and for the cost of shipping a particular dollar's worth of good from a particular region. This will be explained in more detail.

For each type of good, as represented by their industry classifications, we have an estimate of the supply produced in each region and we make estimates of the demand for that good in each region. We also estimate the transportation costs of shipping a dollar's worth of each good from one region to all of the others. Thus, we have the typical linear programming transportation problem.

The transportation problem is solved for each industry separately and two shadow prices are produced for each region. We can interpret the shadow prices as follows: Suppose region j were to demand an additional dollar's worth of the good. Then one shadow price expresses the cost of getting this additional dollar's worth to region j. On the other hand, suppose region j supplied the good. The other shadow price would be the cost of supplying an additional dollar's worth of the good from region j. For example, when trying to explain the location of the furniture industry, we note that one of the principal inputs is lumber. The transportation problem, solved for the lumber industry, would give us the marginal cost of getting an additional dollar's worth of lumber to each of the regions producing furniture. These marginal costs then would be used as one variable in explaining the location of the furniture industry. In addition, the cost of shipping furniture to its markets may also influence the location of the furniture industry. From the shadow prices produced with the transportation problem solved on the furniture industry, we can get the marginal costs of shipping an additional dollar's worth of furniture from each county. These costs will also be used in explaining the location of the furniture industry.

Labor Costs: Initially prevailing wage rates (annual earnings per worker in the industry being located) in the regions will be used as the cost of attaining an additional unit of labor. Wage rates by themselves, however, do not reflect all of the labor costs; not included are job training and labor migration costs. These other costs are generally classified as agglomeration economies or diseconomies of the labor markets.

Let me explain the example. Suppose a firm has a choice in locating between Washington, D.C., and Clay County, Kentucky. Assume that all factors influencing location are equal at both sites, except for labor costs. We know that wage rates are lower in Clay County, but this does not mean the firm will locate there. Washington may have an abundant supply of the type of skilled labor that the firm requires; therefore, the training and moving cost of labor would be zero. If the firm located in Clay County it would have to bear the expense of training local workers or moving trained workers from other areas.

As a longer range project, we plan to estimate the training and moving costs of getting a marginal unit of labor in a certain occupation to each region.

Other Cost: We expect that regional variations in labor and transportation costs will account for most of the regional variation in value of shipments; however, if necessary, other costs or expenses will be considered. Investment cost per square foot or per unit of output may vary by region, not only because of different construction costs, but because of the different sizes of existing plants. It may be that cost of adding an additional unit of capacity for large plants is lower than for small plants. Certain expenses, such as taxes and public utility charges, may also have regional variation significant enough to influence location.

FOOTNOTES

*The author wishes to thank Professor Clopper Almon and Mr. Frank Hopkins for comments on this paper. This project is supported in part by a grant from the Economic Development Administration, U. S. Department of Commerce.

[1] Curtis C. Harris, Jr., State and County Projections: A Progress Report of the Regional Forecasting Project, Occasional Paper, published by the Bureau of Business and Economic Research, University of Maryland, College Park, Maryland (January 1969).

[2] Instead of one, any value could be used that is constant for all regions.

[3] Both this and alternative methods of estimating marginal transportation costs are being investigated by Frank Hopkins in a doctoral dissertation.