LOCATION DETERMINANTS OF MANUFACTURING INDUSTRY IN RURAL AREAS

Eldon D. Smith, Brady J. Deaton, and David R. Kelch

The spatial distribution of economic activity has been the subject of much theoretical study during the last 150 years [11, 19, 20]. The two-state study which provides primary evidence for this article is, to the authors' knowledge, the first attempt to analyze statistically the determinants of industrial location in rural communities with an explicit objective of more enlightened public action at the local, state, and federal levels.[1]

PROBLEM FORMULATION

Federal and state agencies, voluntary organizations, private developers, and local governments make enormous investments of time, effort, and money to create new employment opportunities designed to increase incomes of rural people. Most of these efforts, in one way or another, have been directed toward manufacturing employment. As of 1970, there were 565 submetropolitan municipalities in Kentucky and Tennessee. Of these, 174 (31%) had acquired or optioned one or more industrial sites. These sites were made available to prospective industrial clients either by nonprofit organizations or local government. Such efforts typically require additional local expenditures to provide some combination of water and sewer services, access roads to sites, low-cost financing, and other public services. In addition, revenues may be foregone as a result of tax concessions effected through facility lease-purchase contracts. All these costs represent investments by taxpayers of the community. A serious research and policy objective, then, is to provide local decisionmakers with systematic evidence of the effectiveness of alternative community actions and supporting state and federal government actions.

Relevant research fits mainly into two categories. The first and most common type explores the economic and social impact on the community of new locations or expansions of manufacturing industry.[2] The second and almost totally neglected category examines the likelihood and probable magnitude of returns on these community investments in terms of new locations or expansions of manufacturing employment and payroll.[3]

The authors address two specific items of information germane to the decisions of communities interested in industrial expansion: (1) the probabilities of attracting one or more "acceptable" plants from the total set of manufacturing industries and (2) the effects of alternative community actions on the probabilities.[4] These items are also relevant to the design of supporting state and federal programs.

The authors are, respectively, Professor of Agricultural Economics, University of Kentucky; Associate Professor of Agricultural Economics, University of Tennessee; and Agricultural Economist, Economic Research Service, USDA.

We acknowledge with gratitude the invaluable assistance of numerous colleagues, especially David Debertin and Charles Sappington. Carl Seigrist provided extensive help with computer analysis. The criticisms of our discussant Sandra Batie and three anonymous reviewers have been helpful in refining the article. We, of course, assume responsibility for its content.

[1]Project S-96, Economic, Institutional and Locational Factors Affecting Location of Manufacturing Industry in SubMetropolitan Areas of Kentucky and Tennessee.

[2]As shown in another paper by the senior author, none of the aggregate income and employment multipliers are relevant to decisions by local community action agencies, many being hybrid analyses which are applicable to neither local communities nor multicounty regions [13]. Perhaps the most complete prior assessment of these studies is presented in [17]. Shaffer's [12] is probably the most relevant of the published studies.

[3]Differences in knowledge of cost-related characteristics of various communities also may affect location decisions and are the basis for community publicity and recruitment activities. Esthetic qualities and other intangibles are also potentially involved. They can be conceptualized as reducing costs of acquiring information and as substituting for direct cost-reducing investment, respectively. Neither are considered explicitly in this study.

[4]The larger study (S-96) of which this study is a part also includes maximum likelihood estimates of the effects of community characteristics and actions on direct employment generated by new plants [9, p. 186]. Space restrictions preclude discussion for it here.

GENERAL CONCEPTUAL FRAMEWORK

The general conceptual framework used in specifying the empirical model is based on conventional location theory. Decisions about the location of manufacturing activity were hypothesized to be based on the criterion of minimum combined production and transportation costs for an assumed given geographic distribution of final demand and immobile resource inputs [8, Ch. 9]. However, application to the particular problem addressed in this study is complicated by the heterogeneity of the set of potential industrial locatees and their diverse cost structures.

Evidence of this diversity is the fact that 18 of the 20 two-digit SICs were represented in the manufacturing plants with 20 or more employees established in Kentucky and Tennessee nonmetro communities during the 1970-1974 period. Moreover, no individual SIC represented more than 20 percent of the plants and only one represented more than 10 percent. Thus, the selection of hypothetically important cost-related variables was unavoidably judgmental, because an almost limitless range of services, physical and institutional resources, and environmental and cultural amenities may affect costs of one or more types of manufacturing industry.

MODEL SPECIFICATION

The two probabilistic dimensions of the decision problem examined suggest the use of the linear probability function (or similar mode of analysis) based on data from the actual experience of rural communities. The linear probability function involves specifying a discrete dependent variable (0,1), indicating a dichotomous phenomenon. In this case zero represents the absence of a plant location and one represents one or more locations in the community. The regression coefficients derived can be interpreted as the change in probability of one or more plant locations in the community associated with a one-unit change in the respective independent variable. The total regression equation can be used to evaluate the probability of location associated with any possible combination of values of the independent variables.

Assume that significant coefficients are derived for both community action variables (which the community can manipulate) and other "fixed" factors such as geographic location, labor supply, and access to interstate highways. In this case, by use of the fitted model, "predicted" probabilities of one or more plant locations per unit of time can be derived for each community. In making these predictions one can assume either (1) direct community actions to recruit industry or (2) no community actions (mere passive acceptance of those industries which decide to locate plants in the area). Thus the model provides a basis whereby a community's current situation can be defined in probabilistic terms. Alternatively, it can be used to evaluate potential improvement in probabilities associated with planned community actions which would change the value of one or more variables in the equation.[5] Although this study is based on plants with 20 or more employees, the model can be respecified for different employment minima, employment size ranges, product classes, etc.

The dichotomous dependent variable is not normally distributed, and hence the error term is heteroscedastic. Thus, although the estimates of the regression coefficients are unbiased, the assumption of normality cannot be fulfilled and the T-test for individual regression coefficients is not strictly valid. This problem was recognized from the outset, and procedures for overcoming the problem of the non-normal error term were explored.[6] However, the present form of the model appears to have high predictive utility.

Selection of Independent Variables

Detailed justification of exclusion or inclusion of each potential variable is impossible in this brief article. However, research evidence of comparative cost effects, reported judgments by management personnel of the importance of individual factors [2], and availability of suitable data were the primary bases of selection. A few comments about a small set of potential variables seem sufficient.

Evaluations of comparative cost effects of differences in general tax rates suggested that this variable was not important in determining manufacturing costs [14, 15]. Legal restrictions on overt tax concessions and their *de facto* incorporation into municipal (revenue) bond financing arrangements preclude

[5] A variety of exogenous factors may change over time and thus affect the actual probabilities. Technological change, employment and growth rates of the national economy, and development of highway systems are among them [2, 3, 5, 6]. Hence, the word "evaluate" rather than "predict" connotes the use of the data as a basis for refining judgment, other relevant data being used as well.

[6] Kmenta [10, pp. 265-266] outlines a procedure for overcoming this problem of non-normality of the error term. However, no standard statistical package contains an appropriate routine for Kmenta's procedures. Colleagues of the authors attempted to build their own computational package, but the large sample size has made estimates based on the entire sample impossible thus far.

separate analysis of tax concessions. However, cost-shifting and cost-reducing effectiveness of revenue bond financing packages as a whole has been strongly suggested by cost analyses [15]. Limitations of local community data on the extent and nature of labor union activity and wages of comparable categories of industrial labor precluded inclusion of these variables.[7] Availability, quality, and ownership of industrial sites by nonprofit entities are obviously cost relevant, as is fire protection rating (insurance costs). Investments in education and the presence of a college or university are hypothetically important as means of reducing personnel costs through manpower development and in-service training, and, as local community service amenities, make the community a more acceptable residential area for company personnel. Suitable proxies or measures of general quality of public and private services are not available [7]. Railroad access was included as a dimension of site quality.[8]

The definition of each variable included in the model is given in Table 1 and comments about the relevancy of some are noted in the interpretive comments.

TABLE 1. VARIABLE DEFINITIONS

PLANT LOCATION—Contractual commitment to build, buy, or lease a plant in the community between January 1, 1970 and December 31, 1973. Refers to plants with 20 or more employees.

1. SITE QUALITY[a]—Refers to the "best" designated industrial site in the particular community and indicates the proportion of all sites in the two-state area on which plants were actually located which are equalled or exceeded by specifications of the particular site in all respects. Specifications included water line diameter, sewer line diameter, land area, and access to rail services at boundary of the site.

2. SITE OWNERSHIP[a]—Ownership or option by a nonprofit organization or arm of government.

3. BOND FINANCING AVAILABLE[a]—Expressed willingness to offer industrial revenue bond financing to suitable clients if desired.

4. EDUCATIONAL EXPENDITURES PER PUPIL[a]—Total expenditures including local, state and federal allocations per student in average daily attendance (total county).

5. LABOR AVAILABILITY[a]—Unemployed plus potential additions to the labor force if wages were similar for a population of similar age, sex ratio, race, and education, expressed as a proportion of the existing official labor force (county). (Taken from Stoll, 1977.)

6. FIRE PROTECTION RATING[a]—1970 municipal.

7. COMMUNITY POPULATION—Municipality population 1970.

8. INTERSTATE HIGHWAY ACCESS—Access within the county to interstate highway or four-lane toll road.

9. MILES TO SMSA—Road miles by best available routing to nearest SMSA.

10. MANUFACTURING EMPLOYMENT—County total 1970.

11. COLLEGE PRESENT OR ABSENT—Four year college or university or two year community college affiliated with a state university.

[a]Modifiable by community action.

[b]Note that this measure is not industry specific. Each site was compared with regard to all four characteristics with all sites on which any type of manufacturing industry had located. Communities with no designated sites were assigned a quality score of zero. Assignment of any score greater than or less than zero lowered the significance level of the regression coefficient and the value of R^2.

[7]Previous studies of opinions of management personnel and individual conversations with plant managers suggest that the dominance and quality of collective bargaining activity is important to location decisions.

[8]For a comprehensive assessment of the available research in this subject and its limitations, see Eldon D. Smith, *Location and Growth of Manufacturing Industry in Rural Areas—A Review of Research*. Mississippi State, MS: Southern Rural Development Center, 1978 (in press).

THE KENTUCKY-TENNESSEE REGIONWIDE MODEL

Table 2 shows results of the OLS fit of a linear model in which plants with 20 employees are used as the lower limit for specifying a location during the four-year period 1970-1973 inclusive.[9] The data are for 565 nonmetro communities, 179 of which had one or more plants established during the period, and a total of 321 plants.

The R^2 (0.37) is, by standards for this type of model, reasonably high and several of the variables are significant by the 10 percent criterion.[10] Moreover, none of the coefficients have signs opposite those hypothesized. Overall, these results are analytically encouraging and appear to have important practical meaning. They suggest that quality sites, preferably controlled by public bodies or nonprofit organizations, are indeed fairly powerful recruiting tools. Industrial revenue bond financing, as prior evaluations of cost effects have indicated, is an effective attractant. Improved fire protection and higher expenditures on public education are also significantly associated with new locations. Of the seven significant variables, only two, college and interstate access, are not modifiable by direct community action.[11]

Manufacturing employment and community population which were included as rough proxies for agglomeration effects are not statistically significant in this model and were not in two previous models in which they were entered separately because of their high intercorrelation (r = 0.80).

Labor availability estimated by Stoll's [12] techniques was not statistically significant.[12] Van Veen's [18] finding that in-commuting tends to be positively associated with Stoll's measure of the unutilized labor supply suggests the possibility that the quality of labor may be inversely associated with the available supply.[13]

TABLE 2 LINEAR PROBABILITY FUNCTION REGRESSION MODEL OF FACTORS AFFECTING LOCATION DURING 1970-73

Variable	B Value	T	Significant[c] at P = ?
Intercept	-0.22490	-2.26	0.0102
Site Quality Score[a]	+0.00401	4.25	0.0001
Site Ownership by Public body or Nonprofit Organization	+0.14530	3.31	0.0001
Interstate Highway Access[a]	+0.06451	1.98	0.0241
College Present or Absent[b]	+0.13907	1.76	0.0392
Bond Financing Available[a,b]	+0.19613	5.25	0.0001
Educational Expenditures per Pupil[a]	+0.00040	2.25	0.0121
Miles to S.M.S.A.	-0.00011	-0.22	0.4148
Fire Protection Rating	+0.03386	2.38	0.0088
Manufacturing Employment (1,000)[a]	+0.00451	0.28	0.3882
Labor Availability[a]	+0.00106	1.02	0.1537
Community Population (1,000)[a]	+0.00013	0.02	0.4924

N = 565; R^2 = 0.37; F Ratio = 29.5

[a]One-tailed test.

[b]See Table 1 for definitions.

[c]Dummy variable.

Practical Meaning of the Regression Coefficients

The ratios of the regression coefficients can be interpreted as estimates of the marginal rates of substitution among all variables (Table 3). They provide for an individual community a partial basis for (1) assessing the feasibility of community actions to compensate for its natural disadvantages and (2) evaluating the cost effectiveness of alternative community actions which modify one or

[9]The model is a slight respecification of one originally reported by Kelch [9]. Population size of community and manufacturing employment are included to indicate that they have been evaluated and found not to be associated significantly with new locations in this general model. They are, of course, highly intercorrelated.

[10]The dichotomous dependent variable, by itself, tends to result in a low R^2 if continuous independent variables are specified. Illustrative is a hypothetical case in which the sole determinant (total explanator) is a continuous variable with a range of 0 to 10 and there is an (unknown) threshold value of 5 at which the dependent variable changes from 0 to 1. Residuals would obviously be large unless (1) the threshold value were known in advance and the independent variable were dichotomized at that value or (2) a very complex functional form were specified. However, it is evident that omitted variables also contribute to a low R^2 in this case. A one-tailed test was applied to all variables in this model.

[11]The significant coefficient for presence or absence of a college is consistent with Beale's [1, p. 954] observations about population trends since 1970.

[12]This technique sums numbers of unemployed persons and potential additions to the labor force into a single measure of numbers of nonworking persons who would be available for work at local wage rates. Potential additions to the labor force are estimated by reference to national participation rate norms for groups of comparable age, sex, education, race, and county wage rates.

[13]This is an untested hypothesis which merits further study. If valid, it has obvious relevance to the problem of developing suitable data services on rural manpower. Note that the hypothesis is not necessarily inconsistent with having a high proportion of industries which are attracted by relatively large supplies of cheap, unskilled labor.

TABLE 3. MARGINAL RATES OF SUBSTITUTION AMONG SIGNIFICANT VARIABLES IN LINEAR REGRESSION MODEL

Variable	Site Quality (Points)	Site Ownership or Option	Financing Availability	Fire Protection Rating (Points)	Educational Expenditure per Pupil (Dollars)	Interstate Highway Access	College (Presence)
Site Quality Score (Points)	-1.00	-36.2	-49.0	-8.4	-0.10	-16.1	-34.7
Site Ownership or Option (Yes or No)		- 1.0	- 1.4	-0.23	- 0.003	- 0.44	- 0.96
Financing Availability (Yes or No)			- 1.0	-0.17	- 0.002	- 0.33	- 0.70
Fire Protection Rating (Points)				-1.0	- 0.012	- 1.91	- 4.12
Expenditure per Pupil in Attendance (Dollars)					-1.0	161.2	-347.8
Interstate Highway or Toll Road Access (Yes or No)						-1.0	- 1.41
College or University Present (Yes or No)							- 1.0

TABLE 4. MEAN ESTIMATED PROBABILITIES FOR RURAL COMMUNITIES WITH AND WITHOUT ANNOUNCED NEW PLANT LOCATIONS, 1974-76

	Kentucky		Tennessee		Total	
	Number of Communities	Mean Probability Estimate	Number of Communities	Mean Probability Estimate	Number of Communities	Mean Probability Estimate
One or More New Announced Locations	44	0.50	95	0.55	139	0.53
No New Announced Locations	250	0.20	176	0.24	426	0.22
Total	294	--	271	--	565	--
Communities with Locations and Probability Values of Less than 0.25	6	--	17	--	23	--

more of the variables included in the model.

Note, for example, that a site with a quality score of 16 will compensate for the disadvantage of being located away from an interstate highway or four-lane toll road. Also, bringing a privately owned average quality site under control of a nonprofit agency or governmental unit will fully compensate for the absence of a college or university, as will the availability of revenue bond financing or improving the fire protection rating by four points. About $160 per pupil educational expenditure would be required to offset the disadvantage of being off the interstate system and about $350 would be needed to offset the disadvantage of not having a local college.

A PREDICTIVE TEST OF THE MODEL

That this model has utility as a predictive tool already has been demonstrated by experience in the area during the post-survey period (Table 4). Mean estimated probability values for communities with new announced locations in the post-survey period 1974-1976 were almost identical in the two states (Kentucky mean $\hat{P} = 0.50$, Tennessee mean $\hat{P} = 0.55$), despite a much higher rate of locations in Tennessee. The communities which had no new announced locations had much lower estimated probability values (Kentucky mean $\hat{P} = 0.20$, Tennessee mean $\hat{P} = 0.24$). Moreover, less than 5 percent of the Kentucky communities and 13 percent of the Tennessee communities with estimated probabilities of less than 0.25 had any announced new locations. Of the communities with probability values above 0.25, 23.0 percent in Kentucky and 55 percent in Tennessee had announced locations.

Significance for Investment Decisions

The importance of this predictive power in local decision making is that it provides a basis for assessing the risk of wasted promotional efforts and investments, especially the risk of nonrecovery of investments in sites and other tangible assets.

Even if the contribution of an activity to improvement in the probability of new locations is the same in all resource situations, the absolute probability of new locations is lower and, accordingly, the risk of wasting resources is higher if the community is disadvantageously situated to being with. Thus, the estimated probability values can be extremely important public information even if the individual coefficients are not precise representations of the effectiveness of specific public actions in greatly diverse situations. To decisionmakers, the absolute level of the estimated probability may be as important, or even more important, than the incremental improvement resulting from the communities' investments. As evidence presented in the analysis of a somewhat different issue will show, communities appear in general to respond to their perception of this relative risk (Table 4).

CONCEPTUAL AND EMPIRICAL PROBLEMS

It was recognized from the outset that the assumptions of the linear probability model might not be absolutely valid in extreme cases, i.e., that the functional relationships might not be independent of the magnitudes of the other variables in the model and might not be linear. Yet models analogous to the various forms of the production function (Cobb-Douglas, transcendental) are not readily adaptable to dichotomized data. Therefore, a linear model was fitted with the expectation of testing *ex post* to determine whether it severely misrepresented the relationships in atypical situations.

Some communities are severely disadvantaged in relation to transportation arteries, access to modern training facilities, supplies of labor of relevant categories, and cultural and environmental amenities. Under these conditions, it is difficult to imagine any industry responding to the conventional inducements. On the other end of the spectrum are communities which appear to "have it all" in terms of locational and natural advantages or infrastructure, often supplied from sources outside the local community. In either case, the marginal contribution of community actions to the probability of additional locations seems likely to be small.

Disaggregation of the data into more homogeneous subregions as a test for specification errors tend to support this conjecture. Neither the data for the slow growth regions (Appalachian Kentucky and Cumberland Tennessee) nor those for the rapid growth Eastern Tennessee region (high frequency of locations) showed significant associations between community modifiable variables and new locations. Only in the two large regions with highly dispersed and "moderate" frequencies of new locations were responses to the modifiable variables significantly high. All regions are obviously heterogeneous, each containing some rapid growth nodes and some static communities. Thus more complete testing of the validity of the independence and linearity assumptions is needed. A multiphased procedure was used for this purpose.

Preliminary Tests for Specification Bias

Initially, all continuous variables were logged to test for possible unidirectional curvilinearity. (This step resulted in a small reduction in R^2 values.) Second, residuals from the linear model were plotted against the continuous variables. Third, regressions were compared for two classes of communities defined in terms of access to interstate highways and availability of revenue bond financing.

Residuals from the two-state linear model plotted against the continuous variables[14] provided no evidence of curvilinearities or other patterns suggestive of identifiable specification problems.[15] In addition, the population was separated into two groups for further analysis. The first group was 109 communities which offered revenue bond financing and were located on an interstate highway. The second group was 169 communities off the interstate which did not offer financing. No evidence of interactions was discernible when

[14] Manufacturing employment, labor availability, site quality, and fire insurance rating.

[15] Some heteroscedasticity was noted with regard to residuals plotted against manufacturing employment and labor availability, neither of which are significant. The larger residuals were in the lower ranges of both where the vast majority of observations are found. Biomodality of residuals plotted against site quality has no apparent explanation.

the same regression model, consisting of the remaining variables in the original model, was run on each data set. The resulting regression coefficients and their significance levels followed no consistent pattern.

Classification According to "Natural" Advantages

The two-state linear model was used to classify communities into three groups according to their relative "natural" or situational advantages as determined by variation only in the nonmodifiable variables. Estimated probability values were computed for all communities with the values of modifiable variables held constant at their means. This procedure involves the implicit assumption that variables subject to community control are equal in all cases, and thus communities can be arrayed on the basis of the actual measures of their nonmodifiable characteristics which are included as variables. Separate regression equations allowing variation in all variables subsequently were computed for three groups —the upper 900 observations, the middle 385, and the lower 90 in the probability array (Table 5).[16]

The regression statistics for the middle 385 observations are similar to the general model and lead to the conclusion that any serious errors of estimation would be at the upper and lower extremes.[17] A comparison of the relative magnitudes of the regression coefficients among the three groups (Table 5) provides a partial test of the hypothesis that the marginal contributions of community action variables will be greater in the midrange of communities.

If only the coefficients that were significant at the 0.10 level of t are considered, the regression coefficients in the middle 385 are highest for both site quality and site ownership. The availability of revenue bond financing has the largest coefficient for the high group, with the midgroup second and the low group third. The coefficient for educational expenditure is highest for the low group and third for the

TABLE 5. REGRESSION COEFFICIENTS AND SIGNIFICANCE LEVELS BY ESTIMATED PROBABILITY OF LOCATION GROUPINGS

Variable	Lowest 90 P Values[a]		Middle 385 P Values[b]		Highest 90 P Values[c]	
	B Value	Significant at P = ?	B Value	Significant at P = ?	B Value	Significant at P = ?[d]
Intercept	-0.1797	0.25	0.0047	0.03	-0.0573	0.42
Site Quality Score*	0.0028	0.16	0.0047	0.0001	0.0028	0.16
Site Ownership	0.0803	0.26	0.1689	0.001	0.0427	0.66
Highway Access	--e	--e	0.0561	0.10	-0.1872	0.17
College Present or Absent	--e	--e	--e	--e	0.1286	0.34
Financing Available*	0.1237	0.09	0.1354	0.002	0.3514	0.001
Educational Expenditures per pupil*	0.0007	0.10	0.0004	0.031	0.0006	0.33
Miles to S.M.S.A.	-0.0030	0.04	+0.0005	0.18	-0.0009	0.58
Fire Protection Rating*	0.0080	0.42	0.0233	0.073	0.0571	0.16
Manufacturing Employment	0.0001	0.13	0.0001	0.004	0.00001	0.39
Labor Availability	+0.0225	0.02	-0.0006	0.39	+0.0025	0.36
	R^2 = .22 F =2.3		R^2 = 0.36 F =20.9		R^2 = 0.43 F = 6.03	

[A]P = 0.216-0.225.

[b]P = 0.225-0.315.

[c]P = 0.315-0.499.

[d]One tailed test.

[e]No positive values for this variable.

*Modifiable variables.

[16]Earlier, the hypothesis was advanced that the marginal contributions of community action variables were expected to be low at the lower and upper extremes of the distribution. The number of observations (90) at each end of the spectrum was chosen to represent the extremes and to allow sufficient degrees of freedom to test the hypothesis.

[17]Another analytical result further supported this conclusion. A set of predicted values were calculated with an alternative model which included 20 multiplicative terms designated to capture hypothesized interactions. The simple γ between these predicted values and those from the linear model reported above was 0.936.

midgroup (insignificant for the high group). Fire protection rating is statistically significant only for the midgroup (though second in magnitude). Therefore the hypothesis is supported for three variables directly related to the firm's cost curve, (1) site quality, (2) site ownership, and (3) fire protection rating. Another direct cost variable, revenue bond financing, is more important for the most advantaged communities, though statistically significant across all three levels of "natural" resource endowment.

Community investments in modifiable variables which were *lower* for both the most advantaged and the most disadvantaged communities than for the intermediate group would be consistent with the hypothesized lower marginal contribution at the extremes of the distribution. However, with the exception of educational expenditures per pupil, the level of investment in modifiable variables was by far the greatest in the most advantaged communities (highest $\hat{P}$ values). Also, three categories of investment were slightly higher in the *least* advantaged group of communities (lowest $\hat{P}$ values, Table 6), and none was significantly lower, despite slightly *lower* absolute frequencies of location.

That community decision are guided by considerations of riskiness, i.e., the risk of nonrecovery of tangible asset investments, is strongly suggested by comparison of the most advantaged communities with the other two categories. Though the evidence that increases in modifiable factors make new locations more likely in the most advantaged group is weak at best, very high levels of site quality, more frequent site ownership, and much higher fire protection rating are noted in the most advantageously situated group. Because revenue bond financing is relatively costless to the community, its higher frequency probably reflects a higher level of promotional effort, a somewhat intangible investment.

That the judgments of these more advantaged communities have been valid is reflected in the level of manufacturing employment growth. It was four times as great as that in the intermediate level group, and six times as great as that in the most disadvantaged group; thus a much higher rate of asset recovery is indicated.

Though one cannot rule out some degree of interdependence among variables and some curvilinearity in the relationships, the tests provide no indication that they result in

TABLE 6. MEANS AND STANDARD DEVIATIONS OF OBSERVATIONS GROUPED ACCORDING TO ESTIMATED PROBABILITIES

Variable	Lowest 90 P Values[a]		Middle 385 P Values[b]		Highest 90 P Values[c]	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
P (Proportion with one or more locations)	.22	.004	.24	.26	.38	.06
Site Quality Score*	16	.23	17	23	39	31
Site Ownership*	.31	.46	.25	.43	.54	.50
Highway Access	d	d	.51	.50	.87	.34
College Present or Absent	d	d	d	d	.48	.50
Bond Financing Available*	.39	.46	.32	.44	.60	.47
Educational Expenditures per Pupil*	471	.90	418	103	413	119
Miles to S.M.S.A.	81	31	64	33	59	29
Fire Protection Rating*	2.4	1.4	2.4	1.4	3.6	1.6
Manufacturing Employment	380	573	527	1,051	2,161	3,439
Labor Availability	6	4.4	17	15.8	26	21

[a]P = 0.216-0.225.

[b]P = 0.225-0.317.

[c]P = 0.317-0.499.

[d]No positive values of the variable.

*Modifiable.

serious distortions in the relationships when a linear additive model of the type shown in Table 2 is fitted to the data.

The range in the $\hat{P}$ values was very small when only nonmodifiable variable differences were considered. Eighty-four percent were between $\hat{P} = 0.22$ and 0.35 and the distribution was severely skewed toward the origin. Hence, it seems to be technically possible to compensate for most locational and situational disadvantages. In fact, in every category some communities appear to have rather fully compensated for them. When both modifiable and nonmodifiable variables were allowed to take their actual values in a second regression run, a few of the most disadvantaged communities had predicted values near the mean predicted values of the most advantageously situated group.

Adequacy of Model Specification—Conclusions

Though not conclusive, the evidence suggests that if specification errors persist they mainly affect the validity of the coefficients in the communities at each end of the resource spectrum. For those more advantageously situated, any tendency to overestimate the effects of investment in sites and site-related factors may not be so serious from a practical standpoint because higher rates of locations are more likely to result in recovery of investments. Such errors may be more serious for more poorly situated communities because their rate of locations has been much lower and the risk of nonrecovery of investments somewhat greater.

Despite the recognized omissions and specification problems, the results appear to provide a substantially improved basis for evaluating industrial development potentials and investment risks for local decisionmakers. Though regression coefficients of the linear model may not be equally valid over a wide range of resource situations, the analysis did not imply any serious biases in this regard.

INTERPRETATIONS RELEVANT TO PUBLIC POLICY

Four concluding observations are supported by this analysis. First, the regression coefficients from the linear model are consistent with the belief that programmed community action can improve the probability of acquiring new industries. Second, the evidence suggests that community investments of appropriate types can overcome in major part natural and situational impediments to industrial expansion. Most educational benefits are not specific to the manufacturing industry clients and therefore are not very cost effective for attracting industry.

Third, neither community population nor manufacturing employment was statistically significant when entered together or separately in alternative regression equations. This finding appears to suggest that the regional growth center hypothesis [4] has little support in the data for communities of submetropolitan size in these two states. The effects of the oft-mentioned agglomerative externalities associated with size of community are not in evidence within the submetropolitan size range. Furthermore, only two other variables which could possibly have accounted for lack of community size effects are significant (fire insurance rating and educational expenditure) and only fire insurance rating is highly correlated with population size ($r = 0.6$).[18]

Some advantages of large community size are obvious. Superior ability to spread investment costs for industrial sites, utilities, and promotional activities may make these actions more fiscally feasible. Moreover, larger communities generally have greater ability to internalize employment and other benefits, as commuting levels are typically higher in small communities. Yet in terms of the effectiveness of local actions to attract industry, the smaller community seems to be at no inherent disadvantage in relation to others in the upper range of submetro size communities.

Fourth, any proper policy interpretation of these findings must go well beyond the simplistic notion that "any community can develop that really wants to." The problem of costs and benefits external to the acting community and the difficulties of properly accounting for risk factors for a specific community investment bundle imply that cost-sharing arrangements reflecting spillovers of benefits and costs may lead to more optimal investment patterns. Under such considerations, plant site location could be based largely on access to railroads (a component of site quality) and highways, availability of land for industrial sites, and other cost-related factors. Reliable information on these spillovers is important to any decisions about industrial development strategies and institutional changes required to effectuate them. The information supplied here is but a part of the much broader set of data required.

[18] The view that more frequent locations near interstate highways support the case for agglomeration economies seems flimsy. Interstates provide access to national product and input markets, not just access to specialized services and amenities of metropolitan areas. In any event, the distance between the community of location and the nearest SMSA was not statistically significant.

REFERENCES

[1] Beale, Calvin L. "A Further Look at Nonmetropolitan Population Growth," *American Journal of Agricultural Economics*, Volume 58, December 1976, pp. 953-958.

[2] Carrier, Ronald and Schriver, William R. *Plant Location Analysis: An Investigation of Plant Locations in Tennessee*, Memphis: Memphis State University, 1969.

[3] Garrison, Charles. "Industrial Growth in the Tennessee Valley Region: 1959-1968," *American Journal of Agricultural Economics*, Volume 56, February 1974, pp. 55-60.

[4] Hansen, Niles M. *Intermediate Sized Cities as Growth Centers*, New York: Praeger Publishers, 1971.

[5] Haren, Claude C. "Rural Industrialization in the 1960's," *American Journal of Agricultural Economics*, Volume 52, August 1970, pp. 431-437.

[6] Haren, Claude C. "Location of Industrial Production and Distribution," *in* Larry Whiting, ed., *Rural Industrialization: Problems and Potentials*, Ames: Iowa State University Press, 1974, pp. 3-26.

[7] Hirsch, Verner A. "Output and Costs of Local Government Services," *Proceedings of National Conference on Non-metropolitan Community Services Research*, Washington, D.C.: U.S. Senate, Committee on Agriculture, Nutrition and Forestry, 307-319, July 12, 1977.

[8] Isard, Walter. *Methods of Regional Analysis*, Cambridge: The MIT Press, 1960.

[9] Kelch, David. *Industrial Location in the Nonmetropolitan Communities of Kentucky and Tennessee: 1970-1973*, Ph.D. dissertation, University of Kentucky, Department of Agricultural Economics, 1977.

[10] Kmenta, Jan. *Elements of Econometrics*, New York: Macmillan Publishing Company, 1971.

[11] Nishioka, Hisao and Kramme, Gunter. "Location, Conditions, and Factors: An Evaluation of selected Location Surveys," *Land Economics*, Volume 44, May 1973, pp. 192-205.

[12] Shaffer, Ron E. and Luther Tweeten. *Economic Changes from Industrial Development in Eastern Oklahoma*, Stillwater: Oklahoma State University, Agricultural Experiment Station Bulletin, B-715, 1974.

[13] Smith, Eldon D. and Gene F. Summers. "A Review of Research—How New Manufacturing Industry Affects Rural Areas," unpublished manuscript being prepared for Southern Rural Development Center, Mississippi State, MS, 1978.

[14] Stober, William J. *Taxes, Subsidies and Locational Choice*, University of Kentucky, Department of Economics Discussion Paper 36, 1970, 51 pages.

[15] Stober, William J. and Laurence H. Falk. The Effect of Financial Inducements on the Location of Firms, *Southern Economic Journal*, Volume 36, 1969, pp. 25-35.

[16] Stoll, John R. "Rural Manufacturing Plant Employment Growth in Kentucky and Tennessee: The Effects of Community and Regional Characteristics, 1964-1973," Masters thesis, University of Kentucky, Department of Agricultural Economics, 1977.

[17] Summers, Gene F., Jon Minkoff, Sharon Fumes, Edward Beck, and Frank Clements. *Industrial Invasion of Nonmetropolitan America*, New York: Praeger Publishers, 1976.

[18] Van Veen, Johannes. "Employment Impact on Local Residents of Autonomous Expansion of Manufacturing Industry in Kentucky Nonmetropolitan Counties," Masters thesis, University of Kentucky, Department of Agricultural Economics, 1977.

[19] Von Thunen, Johann. *The Isolated State*, Oxford: Reymon Press, 1966.

[20] Weber, A. *Theory of the Location of Industry*, Chicago: University of Chicago Press, 1909.