NATURAL RESOURCES JOURNAL

Volume 17
Issue 2 *Spring 1977*

Spring 1977

Role or Spatial Relationships in Assessing the Social and Economic Impacts of Large-Scale Construction Projects

James A. Chalmers

THE ROLE OF SPATIAL RELATIONSHIPS IN ASSESSING THE SOCIAL AND ECONOMIC IMPACTS OF LARGE-SCALE CONSTRUCTION PROJECTS

JAMES A. CHALMERS*

This paper presents empirical evidence on two questions that must be answered if the social and economic consequences of large-scale construction projects are to be adequately assessed. The questions can be put most simply in the context of an example. Suppose that a construction project is proposed at location X that will employ 1000 workers continuously for a period of three years. Taking account of the local transportation network and the observed distances construction workers are willing to commute daily, assume that a primary impact region boundary is drawn that contains six communities from which workers could commute daily to work on the proposed project. Assume further that all persons in the impact region live in one of the six communities and that all project workers must necessarily live in one of the six communities and thus no new towns are anticipated.

The impact of the proposed project on each of the six communities will depend on characteristics of the project, on the location of the project relative to the communities, and on characteristics of the individual communities. Most importantly, project and community characteristics will interact to determine the number of new residents moving into each community. The new residents will consist of construction workers and their families in addition to any in-migrants attracted by higher wages or better employment opportunities in the communities due to higher levels of economic activity induced by the construction project. This paper considers only the former source of population increase; it is focused on explaining the number of construction workers that will move into each of the six communities.

This issue can most usefully be approached by dividing it into two sub-issues. First, it must be determined how many of the 1000 jobs will be filled by current residents of the six communities. Once the number of local workers has been determined, subtraction from

---

*Associate Professor of Economics, Arizona State University. The data on which this study is based were collected in the Construction Worker Profile Study sponsored by the Old West Regional Commission. The contributions of many colleagues to that study are acknowledged. Appreciation is also expressed to the Old West Commission for permission to reprint some material originally contained in the *Construction Worker Profile: Final Report.*

1000 will yield an estimate of the number of non-local workers that will have to be absorbed by the six communities in total. The second question can then be addressed; namely, how will the non-local workers choose to distribute themselves among the six communities?

The analysis directed to determining the number of local workers a daily commuting region is capable of supplying is referred to as the "source of supply" analysis while that directed to explaining the distribution of non-local workers among existing communities is called the "residential choice" analysis. Together, these two analyses constitute a set of hypotheses that explain the expected influx of construction workers in each of the communities within the daily commuting region of a proposed construction project.

METHOD

The ability of a community to supply local labor to a potential employer will depend on numbers of unemployed workers, on the extent to which workers presently not in the labor force can be induced to join the labor force, and on the extent to which presently employed persons can be induced to leave their present jobs in order to work for a new employer. Economists would traditionally approach this kind of a problem by postulating a model of the local labor market in which the supply of labor is determined by the optimizing behavior of the community's residents and the demand for labor is derived from demand and cost conditions facing local employers. Likewise, there is a professional literature dealing with the residential choice problem. Typically, the problem has arisen in the context of urban studies where the problem is to determine the residential choices of workers relative to their place of employment. Individuals are hypothesized to choose based on their own preferences in light of the amenities associated with different locations and transport costs associated with the journey to work.

This study did not, however, start from these kinds of traditional models of optimizing behavior. The overall objective of the research of which this study is part, was to provide quantitative information that could be used by planners and researchers to help anticipate the consequences of large, energy related construction projects. There is, of course, nothing necessarily inconsistent between this applied objective and the formal model development and testing procedures often followed by economists. Nevertheless, there was one overriding characteristic of the problem faced here that mitigated against the traditional approach. The study region consists predominantly of small places and rural areas for which there is little or no usable secondary data on the relevant demand and supply determining variables that form the core of the formal economic models. The analysis of this paper begins, therefore, by considering the data base on which the inferences of the study will have to be based. Hypotheses are then formulated and tested in light of the available information.

DATA

Social and economic research motivated by the prospect of large scale development of Western energy resources has been seriously handicapped by lack of relevant data for much of the region in which development may occur. Small area data is limited throughout the country but the situation in the West is particularly severe. First, most of the communities are so small that there is little data pertaining to social and economic characteristics tabulated by the Census. Second, counties tend to be very large and use of regularly published county data as a surrogate for place data is not often possible. Finally, for parts of the region in which change is already occurring, the change has been so large relative to the base that the little information that could be derived from historical data is no longer relevant. In response to this situation, the Old West Regional Commission began a research program referred to as The Construction Worker Profile (CWP) Study which was completed in the spring of 1976.[1] The emphasis of the study was on the collection of primary data, and this data is the basis for the analysis reported on in this paper. The CWP actually consisted of three data collection efforts. A household survey was conducted which consisted of door-to-door interviews with 1432 households in nine communities which have been affected, are being affected, or will be affected by large energy related construction projects. A project survey was undertaken which consisted of distributing and collecting short self-administered questionnaires to construction workers at 14 major construction sites in eight western states. The third effort of the study consisted of a more detailed sociological analysis of three of the household survey communities. Only the results from the project survey are used for the source of supply and residential choice analysis reported on here.

The project survey collected information from over half of the 6000 workers employed on the fourteen large projects underway during the summer of 1975 in the states of North Dakota, Montana, Wyoming, Utah, Colorado, New Mexico, and Arizona. Information

1. Further description of study results and a guide to the publications that have resulted from the study are available on request from the Old West Regional Commission, 1730 "K" Street, N.W., Suite 426, Washington, D.C. 20006.

was obtained on workers' characteristics with respect to household composition, place of residence, previous residence, and occupation. These data provide a limited, but nevertheless usable, number of observations on the local/non-local composition of the work force for each of the fourteen projects and on the residential distribution of the non-local workers. Workers were classified as non-local if their present address was not in the same town as it was before they started work on the project. The essence of the distinction, therefore, is whether the new job necessitated or induced a change of residence.[2]

It should be noted that while the project survey data establish the mix of local and non-local workers on each project and the residential choices of the non-local workers, they do not deal with a number of project or community characteristics that may be important in determining local labor supply or residential choice. The focus, therefore, is largely on the spatial relationship between a project and the communities within its daily commuting region. The hypotheses offered deal primarily with the way in which project size, community size, and distance interact to determine local supply and to influence community choice.

SOURCE OF SUPPLY ANALYSIS

A. *The Hypotheses*

The project survey yields a set of observations on the local/non-local composition of the work force on 14 projects which is too small a number of observations to support statistical analysis. The number of local workers on a given project, however, is a function of the interrelationship of the project site to each community within the project's daily commuting region. An alternative way to formulate the problem, therefore, is to examine the number of local workers supplied by community "i" to project "j" (LW_{ij}). This is a more direct way to approach the problem and has the advantage of substantially increasing the number of observations. Four simple hypotheses with respect to the determinants of LW_{ij} are as follows:

Hypothesis 1: the number of local workers supplied by community i to project j (LW_{ij}) will be positively related to the size of community i (POP_i).

Hypothesis 2: the number of local workers supplied by community i to project j will be a positive function of the total number of employees on project j (E_j).

2. Old West Regional Comm., Construction Worker Profile: Final Report, 5-24 (1976).

Hypothesis 3: the number of local workers supplied by community i to project j will be inversely related to the distance between i and j (D_{ij}), and the effect of distance on LW_{ij} may be non-linear.

Hypothesis 4: the larger the other communities within the project's commuting region (i.e., $\sum_{m=1}^{n} POP_m$ for $m \neq j$), the smaller will be LW_{ij}.

Hypothesis 4 reflects the possibility that the number of jobs available to the residents of a particular community may be limited if there are large competing sources of supply within the commuting region. The rationale for the first three hypotheses are self-evident.

The four hypotheses were tested by estimating cross-section regressions using both of the following functional forms designated equations one and two.

$$LW_{ij} = \alpha_0 POP_i^{\alpha_1} D_{ij}^{\alpha_2} E_j^{\alpha_3} \left(\sum_{m=1}^{n} POP_m\right)^{\alpha_4}, \text{ and} \qquad (1)$$

$$LW_{ij} = \beta_0 + \beta_1 POP_i + \beta_2 D_{ij}^{\gamma} + \beta_3 E_j + \beta_4 \left(\sum_{m=1}^{n} POP_m\right). \qquad (2)$$

Equation one (1) is a constant elasticity form and can be estimated by taking logarithm of both sides of the equation. This particular formulation has the advantage that the exponent determining the effect of distance on the supply of local workers (α_2) is an estimable parameter and need not be specified *a priori.* In the linear formulation of equation two (2) on the other hand, only β_2 can be estimated and the exponent γ has to be specified on *a priori* grounds. The actual procedure followed was to experiment with several values of γ and to use overall goodness-of-fit criteria as the basis on which to judge the most nearly appropriate value.

B. The Data

The data used to estimate equations (1) and (2) have been discussed above in general terms but need to be considered in more detail.

LW_{ij} is an estimate of the total number of workers on project j presently residing in community i who also resided in community i prior to working on project j. The population estimate is based on the sample proportion for each project.[3]

POP_i is the 1970 population of community i.[4]

3. Data Source: Construction Worker Profile, *supra* note 2.

4. Data Source: U.S. Bureau of Census, Dep't of Commerce, United States Census (1970).

D_{ij} is the "best route" highway mileage between community i and project j.[5]

E_j is total employment on project j at the time at which workers were interviewed during the summer of 1975.[6]

A total of 131 observations were derived from the project survey data on communities which supplied one or more workers to a project lying within 100 miles by road of the community. The 100 mile radius is a proxy for the maximum daily commuting distance and was set arbitrarily after looking at the geographical distribution of the local workers around the 14 projects. The 1970 U.S. Census estimates of community populations were felt to be as good an indication as could be obtained of "pre-impact" community population.

C. The Results

Equations (1) and (2) specified above were estimated using ordinary least squares for the 131 communities on which the project survey provided observations. The results are given in Table 1. Each of the estimated coefficients has the anticipated sign. In the multiplicative formulation of the model (equation 1), each of the coefficients is significant at the .95 level except for α_4 and the overall explanatory power of the equation is quite good as indicated by the R^2 of .55. It is interesting that the estimated coefficient on distance is less than 1 and also highly significant. The implication is that although distance inhibits the supply of local workers from a community to a project, doubling distance less than doubles the inhibiting effect of distance. For example, suppose there is a project employing 500 workers located 10 miles from community A and 20 miles from community B. Assume both communities have a population of 3000 and that there is a third community of 3000 within daily commuting distance of the project. The multiplicative formulation would predict $LW_A = 30.8$ and $LW_B = 21.5$, so it can be seen that doubling distance from 10-20 diminishes the supply of local workers by about 30 percent.[7]

The coefficients for equation (2), the linear form, were estimated for six different values of γ, including $\gamma = 1, -.5, -1, -1.5, -2$, and -2.5. The overall explanatory power of the linear form is decidedly

5. Data Source: Rand McNally & Co., Rand McNally Road Atlas (1975).
6. Data Source: Construction Worker Survey, *supra* note 2.
7. For example, the equation to be evaluated for community A is:

$$LW_A = (.018)(3000^{.445})(1/10^{.512})(500^{.981})(1/6000^{.119}) = 30.8.$$

TABLE 1

Ordinary Least Squares Estimates of the Coefficients of the Source of Supply Equations[a]

	α_0	α_1	α_2	α_3	α_4	R^2	F	N
	.018	.445	−.512	.981	−.119	.55	38.56	131
	(−3.3)	(9.9)	(−6.9)	(6.4)	(−1.2)			
	β_0	β_1	β_2	β_3	β_4			
γ = 1.0		.002	−.377	.024	−.000	.31	13.84	131
		(6.3)	(−4.7)	(2.2)	(−1.6)			
γ = − .5	−6.054	.002	36.846	.027	−.000	.22	8.98	131
	(−.8)	(5.3)	(2.4)	(2.3)	(−1.3)			
γ = −1.0	.507	.002	26.671	.026	−.000	.20	7.85	131
	(.1)	(5.0)	(1.5)	(2.2)	(−1.3)			
γ = −1.5	2.018	.002	19.335	.026	−.000	.19	7.41	131
	(.3)	(4.9)	(.9)	(2.2)	(−1.3)			
γ = −2.0	2.531	.002	13.366	.026	−.000	.19	7.25	131
	(.4)	(4.9)	(.6)	(2.2)	(−1.4)			
γ = −2.5	2.746	.002	8.721	.026	−.000	.19	7.19	131
	(.4)	(4.9)	(.3)	(2.2)	(−1.4)			

[a]Numbers in parentheses below the estimated coefficients are calculated "t" values. With $N \geq 120$, the null hypothesis that the coefficients equal zero can be rejected on a one-tail test at the .95 level if the calculated "t" ≥ 1.658.

higher when γ assumes the value of 1 than is the case when the reciprocal of distance is used raised to powers from .5 to 2.5.

Despite the statistical significance of the coefficients and the confirmation of the four simple hypotheses offered earlier in this section, the motivation for this work is not so much to confirm the obvious fact that larger communities will be able to supply more construction workers than smaller communities, but to see whether the simple relationships between community size, project size, and distance appear to have predictive capability in the impact assessment process.

Table 2 compares the results obtained when the number of local workers is estimated for each of the 14 sample projects using three different procedures. The first method is to use the regression results obtained in the multiplicative form to estimate the number of local workers, the second is to use the regression results from the linear model with $\gamma = 1$, and the third, in order to provide a standard of comparison, is to project the number of local workers under the naive hypothesis that each project would have the same proportion of local workers as did all 14 projects in the aggregate.

The table shows that all three methods result in sizable errors and that no method gives consistently smaller projection errors than the other. In fact, the multiplicative model has the smallest error for four projects, the linear model for five projects, and the naive model for five projects. Another way in which the three methods can be evaluated is by comparing the mean squared error (MSE) obtained under each. This criterion gives increasingly heavy weight to large errors which is appropriate since the costs of being wrong are likely to increase more than in proportion to the size of the error. The MSE for the multiplicative model with $\gamma = 1$ is 5802.4, for the linear model is 3036.3, and for the naive model is 6589.7. Thus, although each of the models makes small errors for about the same proportion of the projects, the linear model has the decided advantage of making relatively fewer large errors than either the naive or the multiplicative models.

Confidence in the linear model is further increased because several of the cases of large errors can be rationalized on the basis of additional information on specific projects. For example, there are large over estimates of the supply of local labor to Coronado, Colstrip, Emery, Huntington, Texas Gulf, and Wyodak. The contractor at both Emery and Huntington was JELCO, and the Construction Worker Profile study team was led to believe that JELCO is unique among the general contractors in the region in maintaining a salaried, 12-month work force which it is capable of moving as necessary. If

TABLE 2

LOCAL WORKER PROJECTIONS UNDER THREE VARIATIONS OF THE SOURCE OF SUPPLY MODEL

		Estimated Local Workers			*Absolute Error*		
Project Number and Name	*Actual Local Workers*	*Multiplicative Model*	*Linear Model with γ=1*	*Naive Model*	*Multiplicative Model*	*Linear Model with γ=1*	*Naive Model*
1 Coronado 1, 2, 3	45.0	32.9	65.0	50.9	12.1	–20.0	–5.9
2 Craig 1, 2–Yampa Power Plant	138.1	128.0	155.9	163.5	10.1	–17.8	–25.4
3 Hayden 2	179.7	139.9	184.7	248.9	39.8	–5.0	–69.2
4 Colstrip 1, 2	118.0	120.2	185.9	250.6	–2.2	–67.9	–132.6
5 Center–Milton R. Young	249.9	209.9	208.8	169.8	40.0	41.1	80.1
6 Leland Olds	268.8	193.1	234.3	199.0	75.7	34.5	69.8
7 San Juan 1	387.2	160.3	236.3	165.5	226.9	150.9	221.7
8 Emery	93.0	67.4	146.6	78.4	25.6	–53.6	14.6
9 Huntington 2	148.7	133.3	217.7	146.7	15.4	–69.0	2.0
10 Jim Bridger 2, 3	224.9	113.0	214.2	280.7	111.9	10.7	–55.8
11 Texaco Lake Expansion	101.4	49.4	117.0	100.5	52.0	–15.6	.9
12 Sun Oil–Cordero Mine	70.2	21.6	37.0	66.7	48.6	33.2	3.5
13 Texas Gulf Sulphur	86.9	62.0	121.9	125.6	24.9	–35.0	–38.7
14–Wyodak	4.4	43.0	37.6	67.0	–38.5	–33.2	–62.6

this is so, it explains the lower than expected penetration of local workers on the JELCO projects. Texas Gulf Sulphur also had low local labor involvement but this may very well be due to the fact that at the time when Texas Gulf Sulphur began construction in August of 1973, there were already 3000 employees working on the Jim Bridger units only 80 miles away. Likewise, the local worker involvement on the Coronado project in Arizona may have been reduced because of the simultaneous activity on the Cholla project, less than 100 miles away. Possible explanation for the smaller than expected number of local workers on Colstrip and Wyodak is not as clear.

If, as is often the case, the social and economic implications of a construction project which was not scheduled to begin for 4-6 years were being assessed, nothing would be known about local labor market conditions at the time construction is to begin, about the identity of the contractor, or about a number of other community and project-specific pieces of information. In this circumstance, a judgment about the number of in-migrating construction workers to be associated with the project would have to be based on an estimate of the number of local workers supplied to the project and this would seem to be best derived using the regression results obtained with the linear model with $\gamma = 1$. As the starting date of construction approached, however, the local worker projection would have to be continually evaluated in light of better information as it became available.

RESIDENTIAL CHOICE ANALYSIS

Once an estimate has been made of the number of non-local workers expected on a project, the next step is to estimate where the workers will choose to live within the commuting region. Will their residential choices reflect preferences with respect to community size, and does distance from the project site affect choice of community? Are there other variables that seem to influence community choices?

A. The Hypotheses

The residential choice model is based on the idea that the relative attractiveness of communities can be measured by observing the relative numbers of non-local in-migrants from a given project who settled in each. Specifically,

$$NLW_i = (A_i/A)TNLW_j, \tag{3}$$

where

NLW_i is the number of non-local workers settling in the community i,

$TNLW_j$ is the total number of non-local workers on project j,

A_i is the attractiveness of the *ith* community, and

A is the sum of A_i over all communities.

The basic hypotheses of the model center on the determinants of community attractiveness. Equation (4) suggests that the attractiveness of the *ith* community as a place of residence for non-local workers from the *jth* project is related by a multiplicative constant to the size of the community (POP_i) and on the distance separating the community and the project (D_{ij}).

$$A_i = k(POP_i^{\beta}/D_{ij}^{\gamma}) \qquad (4)$$

This is obviously a simple hypothesis, and its rationale is discussed more critically below. First, however, it is necessary to see how the hypotheses in equations (3) and (4) can be tested and how the effect of distance (i.e., the value of γ) and of size (i.e., the value of β) can be estimated.

The assumption is made that the number of non-local residents in community 1 (NLW_1) relative to the number in community 2 (NLW_2) is a reflection of the attractiveness of 1 (A_1) relative to 2 (A_2). That is,

$$NLW_1/NLW_2 = A_1/A_2,$$

or in general, for any pair of communities s and t,

$$NLW_s/NLW_t = A_s/A_t \quad \text{for } s,t = 1,2,\ldots n,$$

where n is the number of communities in the commuting region of project j. Substituting equation (4) for A_s and A_t,

$$NLW_s/NLW_t = \left(\frac{POP_s^{\beta}/D_{sj}^{\gamma}}{POP_t^{\beta}/D_{tj}^{\gamma}}\right). \qquad (5)$$

Ordinary least squares can then be used to estimate the size and distance elasticities once equation (5) is made linear by taking a logarithmic transformation.

$$\log NLW_s - \log NLW_t = \beta(\log POP_s - \log POP_t) - \gamma(\log D_{sj} - \log D_{tj}). \qquad (6)$$

Every pair of communities within the commuting region of a project provides an observatioon relative attractiveness, and by using all communities which had at least three NLW residents, the project survey provided a total of 140 such observations as the basis on which the hypotheses behind (6) can be tested.

The validity of the model depends on the assumptions that the attractiveness of a community increases with its size, at least over the range of sizes considered in this study, and, equally important, that the ability and speed with which a community can expand to absorb new residents is also a function of size. Support for this contention comes from considerations like the ease of generating credit for purposes of residential construction or the level of sophistication of the managers of municipal services. The effect of distance is to reduce desirability, but, as with the source of supply model, the important question concerns the quantitative magnitude of the effect.

B. The Results

If the non-local workers on a project distribute themselves between two communities, then there is only a single observation on relative attractiveness (NLW_1/NLW_2). If there are three communities, however, there are three observations (NLW_1/NLW_2, NLW_1/NLW_3, NLW_2/NLW_3), and for four communities there are six observations. In general, where N is the number of communities, the number of observations in the regression model equals N(N-1)/2. For the fourteen projects, the regression result obtained using ordinary least squares is shown below in equation (7),

$$NLW_s/NLW_t = \left(\frac{POP_s^{.385}/D_{sj}^{.638}}{POP_t^{.385}/D_{tj}^{.638}}\right) \qquad (7)$$

and both estimated coefficients have t-ratios in excess of 7.

The absolute size of the distance coefficient is consistent with the evidence from the source of supply analysis. Contrary to a strict gravity model formulation in which distance has an inverse effect proportional to its square; i.e., 20 miles has an inhibiting effect relative to 10 miles as 4:1, the results here suggest that the distance effect may have an exponent less than 1. For an example, using the .638 exponent, 20 miles would have an inhibiting influence relative to 10 miles as 1.6:1. This result is not inconsistent with the casual observation of researchers on the attitude of Western workers with respect to commuting substantial distances.

Examination of the residuals from equation (7) revealed that they were not randomly distributed by project. This is not unexpected and indicates that some variables specific to the region are acting systematically to affect community choice. For example, an initially small community may have been quicker to provide housing than would be expected on the basis of its size, and therefore ended up with many of the non-local workers as residents even though there may have been larger communities in the near vicinity.

The extent of the regional variation in the explanatory power of the model was investigated by re-estimating the model for geographic subregions. There were four areas, each of which included a pair of projects, for which there was a sufficient number of observations to re-estimate the model. The results are in Table 3. It can be seen that although there is variation in the size of the coefficients from area to area, there is additional evidence for the result that γ lies between .5 and 1. The elasticity of attractiveness with respect to community size varies more. For the Wyoming and North Dakota areas the elasticity is about .25 while for the Utah area it is about .80 and for Northwest Colorado about 1.2.

TABLE 3

SUBGROUP REGRESSIONS FOR COMMUNITY CHOICE MODEL[a]

	β	γ	n
All Projects	.385 (7.2)	.638 (7.4)	140
Craig-Hayden	1.227 (18.6)	.522 (8.4)	13
Center-Leland Olds	.283 (5.3)	.765 (7.4)	38
Emery-Huntington	.795 (4.0)	1.183 (4.3)	20
Jim Bridger-Texas Gulf	.263 (2.4)	.051 (.2)	42

[a]Calculated t-ratios are in parentheses below the estimated coefficients.

These results suggest that although the hypotheses underlying the residential choice model are supported by the data, the variation of the estimated parameters from area to area indicates that region-specific characteristics such as provision of housing by construction companies, the nature of transportation options, availability of mortgage financing, etc. have to be taken into account before much progress can be made in predicting the communities where non-local construction workers will choose to live.

SUMMARY AND CONCLUSIONS

The purpose of this paper was to report on efforts to find regularities in the spatial relationships of large construction projects to the communities that surround them. The particular question investigated concerned the number of non-local workers that could be expected as the result of a project and the way in which they would distribute themselves among the communities within the commuting region of the project. This question was pursued with the explicit recognition of the fact that many circumstances, particular to the project and the communities at the time at which construction commenced, played a role in determining the required number and dispersion of non-local workers. The question was whether these considerations were paramount or whether generalizations could be made on the basis of community size, project size, and the spatial relationship of the project to the communities that surround it that would be useful in the impact assessment process.

The results indicate that although the hypothesized interactions of size and distance are easy to identify in the data, the estimates are likely to be subject to errors and should only be used as a starting point in the impact assessment process. For certain kinds of purposes, however, such as in regional or programmatic studies where large numbers of projects are being assessed, or where little is known about a proposed project except its location and size, the relationships reported here provide a basis for making initial conjectures about the nature of community impacts likely to be associated with the project(s). The impression must not be given, however, that this is a substitute for detailed consideration of other project and community characteristics that will ultimately determine the social and economic implications of a large construction project. As research is able to be focused on a particular project and as the date on which construction begins approaches, the general models outlined in this paper become less relevant relative to the specifics of the particular situation.