JOURNAL OF REGIONAL SCIENCE, VOL. 9, NO. 2, 1969

AN ECONOMIC DERIVATION OF THE "GRAVITY LAW" OF SPATIAL INTERACTION*

J. H. Niedercorn and B. V. Bechdolt, Jr.†

Spatial interaction models have been based upon analogies to physical laws of Newtonian physics. These models have questionable value in theoretical explanations of people's interaction in space. The purpose of this paper is to derive the "gravity law" of spatial interaction within the framework of utility theory, thereby providing a theoretical explanation for well-known empirical regularities. Section 1 of this paper provides a brief background in spatial interaction models, and Section 2 presents the economic derivation of the "gravity law".

1. SPATIAL INTERACTION MODELS

Spatial aspects of economic phenomena have been virtually excluded from the mainstream of orthodox economic theory. Writings of location theorists, such as Von Thünen [33], Weber [35], Lösch [20], Hoover [14], Dunn [11] and Isard [17], have essentially remained separate from general price and distribution theory, which has abstracted from space.

The interaction of persons and things in space has been studied by economists, demographers, geographers, sociologists, planners and others. Models used by these investigators were developed from the writings of location, land, and rent theorists, most of them originally through analogies to physical relationships and were called social physics.

As used in the behavioral sciences, spatial interaction models describe social phenomena in space, such as population migration, flow of goods, money, and information, traffic movement and tourist travel. These models are used extensively in market area, transportation, tourism and demographic studies.

Many factors affect the movement or interaction of persons and things in space. If social phenomena are conceived as occurring between geographic areas or points, each of these factors can be categorized as (1) an origin factor, (2) a destination factor, or (3) a linkage factor. Origin factors are characteristics of an origin that

* The authors are grateful to Gerhard Tintner for his suggestions and comments.

† The authors are Associate Professor of Economics and Urban and Regional Planning at the University of Southern California and Assistant Professor of Economics and Statistics at California State College, Los Angeles, respectively.

Date received: September, 1968.

usually indicate the capacity of an origin to interact with all possible destinations. Destination factors are characteristics of a destination that account for the relative magnitude of the interaction of the destination with all origins. Often, destination factors reflect the attractiveness or appeal of a destination for persons in all origins. Linkage factors are characteristics of the relationships between a particular origin and destination.

Because spatial interaction models are quantitative, they are usually expressed in the form of mathematical equations. The most important origin, destination, and linkage factors are made origin, destination, and linkage variables in mathematical functional relationships. Typically, the dependent variable is a quantitative measure of the spatial phenomenon being studied and is specified to be a function of one or more origin, destination, and linkage variables.

There are two principal types of spatial interaction models—gravity models and potential models. Gravity models are used to calculate the number of interactions between two geographic areas or points. Potential models are used to calculate an index of potential interaction at an area or point.

Spatial interaction models can be classified according to (1) the type of data used, (2) the type of interaction being studied, and (3) the point of view from which the interaction is being studied.

Type of Interaction	Type of Data	
	Cross-sectional	Time Series
Unidirectional	Origin-centric Models Single-origin/ Multiple-destination Multiple-origin/ Multiple-destination Destination-centric Models Single-destination/ Multiple-origin Multiple-destination/ Multiple-origin	Single-link Model
Bidirectional	Multiple-link Model	Single-link Model

The fundamental idea underlying spatial interaction models is that the degree of interaction between two geographic areas is a function of (1) the degrees of concentration of persons or things in the two areas and (2) a measure of the distance separating the two areas. This fundamental idea is analogous to, and was originally derived from, the Newtonian laws of force and energy. Gravity and potential models evolved as a part of the early work of social physicists who believed that social phenomena could be explained by physical laws.

The earliest statement of human interaction was made by Carey [5] in 1858. Carey believed that man is to society as a molecule is to matter. He stated that the more persons concentrated in a given area, the more attractive force there is exerted by that area. Carey defined the "gravity law" of spatial interaction by saying that the degree of attraction varies directly with the mass, or concentration of persons or things, and inversely with distance. Later writers developed, expanded, modified and applied these concepts; these writers include Ravenstein [23], Young [36], Reilly [24], Bossard [4], Stewart [25] and [31], Zipf [37] and [38], Stouffer [32], Converse [9], Dodd [10], Anderson [1] and [2], Carrothers [7], Voorhees [34], Isard and Freutel [19], Harris [13], Carroll [6], and Huff [15] and [16].

2. "GRAVITY LAW" OF SPATIAL INTERACTION

The empirical regularity of the "gravity law" of spatial interaction has tended to make it a widely used tool of spatial analysis. But as an analogy, it suffers from the lack of a firm theoretical foundation. The following derivation of the "gravity law" provides this foundation within the framework of utility theory. The derivation is given in terms of trips made by persons from a single origin to many destinations (single-origin/multiple-destination model).

2.1. *General Derivation*

Assume that a defined subnational region is composed of $n + 1$ areas. The initial task is to study the tripmaking behavior of an individual, k, at origin area $i(i = 0)$ with each destination area $j(j = 1, 2, \cdots, n)$. Assume that there is only one person or thing at each destination, j, with which the individual at origin i would like to interact per trip. Then this individual's utility of tripmaking (net of any disutilities of tripmaking) from origin i to destination j provides a first approximation of the individual's utility of interaction with destination j.

$$(1) \qquad {}_kU'_{ij} = f({}_kT_{ij})$$

where

${}_kU'_{ij}$ = net utility of individual k at origin i of interacting with persons or things at destination j, per unit time, and

${}_kT_{ij}$ = number of trips taken by individual k from origin i to destination j, per unit time.

An individual's total net utility of interaction with all destinations under the above assumptions is

$$(2) \qquad {}_kU'_i = \sum_{j=1}^{n} f({}_kT_{ij})$$

where

${}_kU'_i$ = total net utility of individual k at origin i of interacting with persons and things at all destinations, per unit time.

However, in reality, there is usually more than one individual or thing at each destination with which the individual at origin i would like to interact. As a rough approximation, the number of such persons or things at each destination is assumed to be proportional to its population. Consequently, assuming that an individual can make only one interaction per trip, an individual's utility function of interaction with persons or things at all destinations is

$$(3) \qquad {}_kU_i = a \sum_{j=1}^{n} P_j f({}_kT_{ij})$$

where

$_kU_i$ = total net utility of individual k at origin i of interacting with persons or things at all destinations, per unit time,

P_j = population of destination j, and

a = constant of proportionality.

However, the individual cannot take an unlimited number of trips (assuming $_kU_i > 0$ for all $_kT_{ij} > 0$), thereby increasing his utility without limit. He is constrained by both (1) the total amount of money that he is willing to spend on travel out of a limited income and (2) the total amount of time that he is willing to allocate for travel. The monetary constraint requires simply that the total amount spent for travel to all destinations must be equal to or less than the total amount budgeted for travel.

$$(4) \qquad {}_kM_i \geq r \sum_{j=1}^{n} d_{ij}\, {}_kT_{ij}$$

where

$_kM_i$ = total amount of money individual k located at origin i is willing to spend on travel, per unit time,

r = cost per mile of distance travelled, and

d_{ij} = distance between origin i and destination j.

Correspondingly, the time constraint requires that the total amount of time spent travelling to all destinations must be equal to or less than the total amount of time allocated for travel.

$$(5) \qquad {}_kH_i \geq 1/s \sum_{j=1}^{n} d_{ij}\, {}_kT_{ij}$$

where

$_kH_i$ = total amount of time individual k located at origin i is willing to allocate for travel, per unit time and

s = average speed at which people in the region travel.

For any given individual, k, one of these two constraints will be more binding than the other. The relevant constraints for individual k can be found by comparing the maximum distance that can be travelled in the region, $_kM_i/r$ or $s\, _kH_i$: if $_kM_i/r < s\, _kH_i$ money is the binding constraint; if $_kM_i/r > s\, _kH_i$ time is the binding constraint.

The total net utility of individual k for interaction with persons or things at all destinations when constrained by money is obtained by maximizing[1]

$$(6) \qquad {}_kU_i^* = a \sum_{j=1}^{n} P_j f({}_kT_{ij}) - \lambda\left(r \sum_{j=1}^{n} d_{ij}\, {}_kT_{ij} - {}_kM_i\right)$$

where

$$\lambda = \text{Lagrangian multiplier}$$

with respect to the number of trips to each destination per unit time. The first order

[1] A parallel derivation can be made using the time constraint simply by substituting $_kH_i$ for $_kM_i$ and s for r.

conditions for maximizing Equation (6) are

$$
\begin{aligned}
&\frac{\partial\, {}_kU_i^*}{\partial\, {}_kT_{i1}} = aP_1 \frac{\partial f({}_kT_{i1})}{\partial\, {}_kT_{i1}} - \lambda r d_{i1} = 0 \\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad \vdots \\
&\frac{\partial\, {}_kU_i^*}{\partial\, {}_kT_{in}} = aP_n \frac{\partial f({}_kT_{in})}{\partial\, {}_kT_{in}} - \lambda r d_{in} = 0 \\
&\frac{\partial_k U_i^*}{\partial \lambda} = r \sum_{j=1}^{n} d_{ij}\, {}_kT_{ij} - {}_kM_i = 0
\end{aligned} \tag{7}
$$

Eliminating λr from the first n partial derivatives in Equation (7) yields the following $n - 1$ equations plus the partial derivative with respect to λ.

$$
\begin{aligned}
&P_2 \frac{\partial f({}_kT_{i2})}{\partial\, {}_kT_{i2}} - \frac{d_{i2}}{d_{i1}} P_1 \frac{\partial f({}_kT_{i1})}{\partial\, {}_kT_{i1}} = 0 \\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&P_n \frac{\partial f({}_kT_{in})}{\partial\, {}_kT_{in}} - \frac{d_{in}}{d_{i1}} P_1 \frac{\partial f({}_kT_{i1})}{\partial\, {}_kT_{i1}} = 0 \\
&r \sum_{j=1}^{n} d_{ij}\, {}_kT_{ij} - {}_kM_i = 0
\end{aligned} \tag{8}
$$

The utility maximizing values of ${}_kT_{ij}$ for all j can be found by solving Equations (8) simultaneously. However, an explicit solution can be found only after the utility of tripmaking function, $f({}_kT_{ij})$, is specified. In the following sections, solutions to Equations (8) are derived using a logarithmic and a power function of tripmaking. For each solution, assume that each individual at origin i has the same form of utility function.

2.2. *Solution with a Logarithmic Utility of Tripmaking Function*

If

$$f({}_kT_{ij}) = \ln {}_kT_{ij} \tag{9}$$

then by substitution of $\partial f({}_kT_{ij})/\partial\, {}_kT_{ij} = 1/{}_kT_{ij}$ into Equations (8), it is easy to derive

$$ {}_kT_{ij} = \left(\frac{{}_kM_i}{r}\right)\left(\frac{P_j}{\sum_{j=1}^{n} P_j}\right)\left(\frac{1}{d_{ij}}\right). \tag{10}$$

Equation (10) shows that the utility maximizing number of trips taken by individual k to destination j per unit time is:

(1) directly proportional to total distance travelled by individual k to all destinations per unit time, ${}_kM_i/r$,

(2) directly proportional to the fraction of the region's population that is located at destination j, $P_j/\sum_{j=1}^{n} P_j$, and

(3) inversely proportional to the distance from origin i to destination j.

The total number of trips taken by all individuals from origin i to a particular destination, j, is obtained by summing the trips from origin i to destination j taken by the m individuals at origin i to obtain

$$\begin{aligned} T_{ij} &= \sum_{k=1}^{m} {}_kT_{ij} \\ &= \left(\frac{1}{r}\right)\left(\frac{1}{d_{ij}}\right)\left(\frac{P_j}{\sum_{j=1}^{n} P_j}\right)\left(\sum_{k=1}^{m} {}_kM_i\right) \\ &= \left(\frac{M_i}{r}\right)\left(\frac{P_j}{\sum_{j=1}^{n} P_j}\right)\left(\frac{1}{d_{ij}}\right) \end{aligned} \tag{11}$$

where

T_{ij} = total number of trips taken by all individuals from origin i to destination j, per unit time and

M_i = total amount of money that all individuals at origin i are willing to spend for travel to all destinations, per unit time.

If travel time, rather than money, is the relevant constraint, then M_i/r in Equation (11) is simply replaced by sH_i, where H_i is total time allocated by all individuals at origin i for travel to all destinations, per unit time. The equations for T_{ij}, using either the money or the time constraint, are similar to the version of the "gravity law" showing that T_{ij} is inversely proportional to d_{ij} raised to the first power.

2.3. *Solution with a Power Utility of Tripmaking Function*

If

$$f({}_kT_{ij}) = {}_kT_{ij}^{b} \qquad (0 < b < 1) \tag{12}$$

then by the substitution of

$$\frac{\partial f({}_kT_{ij})}{\partial\, {}_kT_{ij}} = b\, {}_kT_{ij}^{b-1}$$

into Equation (8), the utility maximizing number of trips from origin i to destination j made by individual k can be derived[2] as

$${}_kT_{ij} = \left(\frac{{}_kM_i}{r}\right)\left(\frac{P_j^{[1/(1-b)]}}{\sum_{j=1}^{n} \frac{P_j^{[1/(1-b)]}}{d_{ij}^{[b/(1-b)]}}}\right)\left(\frac{1}{d_{ij}^{[1/(1-b)]}}\right). \tag{13}$$

Equation (13) shows that the number of trips taken by individual k from origin i to destination j per unit time is:

(1) directly proportional to the total distance travelled by individual k to all destinations, per unit time, ${}_kM_i/r$,

[2] This derivation is available from the authors on request.

(2) directly proportional to the population of each destination raised to the power $[1/(1 - b)] > 1$,
(3) inversely proportional to the constant,

$$\sum_{j=1}^{n} \frac{P_j^{[1/(1-b)]}}{d_{ij}^{[b/(1-b)]}}$$

and
(4) inversely proportional to the distance between origin i and destination j, raised to the $[1/(1 - b)]$ power.

The total number of trips per unit time taken from origin i to destination j by all individuals at i is

$$T_{ij} = \sum_{k=1}^{m} {}_kT_{ij} = \left(\frac{M_i}{r}\right)\left(\frac{P_j^{[1/(1-b)]}}{\sum_{j=1}^{n} \frac{P_j^{[1/(1-b)]}}{d_{ij}^{[b/(1-b)]}}}\right)\left(\frac{1}{d_{ij}^{[1/(1-b)]}}\right). \tag{14}$$

If travel time is the relevant constraint, then M_i/r in Equation (14) is replaced by sH_i.

It is common, especially when statistical methods such as the method of least squares are used, to find that the population weight, P_j, is raised to an estimated power that is usually different from the power of d_{ij}. The values of P_j in the utility function can be raised to a power c, where $c > 0$. The resulting utility maximizing number of trips taken by individual k to destination j per unit time is[3]

$$ {}_kT_{ij} = \left(\frac{{}_kM_i}{r}\right)\left(\frac{P_j^{[c/(1-b)]}}{\sum_{j=1}^{n} \frac{P_j^{[c/(1-b)]}}{d_{ij}^{[b/(1-b)]}}}\right)\left(\frac{1}{d_{ij}^{[1/(1-b)]}}\right). \tag{15}$$

The total number of trips taken to destination j per unit time by all individuals at origin i is

$$T_{ij} = \sum_{k=1}^{m} {}_kT_{ij} = \left(\frac{M_i}{r}\right)\left(\frac{P_j^{[c/(1-b)]}}{\sum_{j=1}^{n} \frac{P_j^{[c/(1-b)]}}{d_{ij}^{[b/(1-b)]}}}\right)\left(\frac{1}{d_{ij}^{[1/(1-b)]}}\right) \tag{16}$$

If time is the relevant constraint, then M_i/r in Equation (16) is replaced by sH_i

2.4. *Reconciliation of Derivations with the Traditional Gravity Model*

In its most general form, the traditional gravity model is

$$T_{ij} = \frac{\alpha P_i^{\beta} P_j^{\gamma}}{d_{ij}^{\delta}} \tag{17}$$

---

[3] This derivation is available from the authors on request.

where P_i is the population of origin i and α, β, γ, δ are parameters. Equations (11), (14), and (16) are derived solutions for utility maximization of spatial interaction of all individuals at origin i, subject to the monetary budget constraint. These equations can be reconciled with Equation (17) by assuming that M_i is proportional to the population of the origin, P_i, raised to the power β. A realistic alternative reconciliation is to express M_i as

$$\begin{aligned} M_i &= uY_i \\ &= (uP_i)\left(\frac{Y_i}{P_i}\right) \end{aligned} \tag{18}$$

where

Y_i = aggregate money income at origin i and

u = fraction of total money income at the origin that is budgeted for travel.

Here, per capita income at origin i, Y_i/P_i, is a constant, and M_i is proportional to the population of the origin times its per capita income.

If time is used as a constraint, modified versions of Equations (11), (14), and (16) are derived solutions for utility maximization of spatial interaction of all individuals at origin i, subject to the time constraint. Similar reasoning based on an even more likely assumption of proportionality of H_i and P_i leads to the same results.

2.5. *Using the Spatial Interaction Equations*

Depending upon the specific form of the utility function of tripmaking and the relevant constraint, money or time, Equations (11), (14), or (16), or their time-constraint versions, could be used to estimate the number of trips taken from origin i to destination j per unit time by all individuals at origin i. For most middle class people, time is generally the relevant constraint for travel between areas within a metropolitan region, and money is the relevant constraint for travel to areas outside the region. Most poor people in metropolitan regions tend to be limited by the money constraint rather than the time constraint for travel among areas of the region, as well as to areas outside the region. Therefore, spatial interaction models used for forecasting traffic from a particular area should be chosen so as to incorporate the constraint that is most relevant to the persons at the origin.

The spatial interaction models presented in this paper can be fitted econometrically by applying regression analysis to a logarithmic transformation of Equation (17). The models are useful for forecasting since the values of r, s, M_i, H_i and P_j can be estimated independently of the regression analysis, and can be projected using trend values fitted to historical data on the variables. Data from budget studies and travel time studies would be necessary to estimate M_i and H_i, respectively, for each area.

One further application of the models presented in this paper is possible. Starting from an econometrically estimated gravity relationship of the form given in Equation (17), it is possible to make statistical inferences about the form of the utility function. If γ and δ are found to be in the neighborhood of one, an approximately logarithmic utility function can be inferred. However, if δ is significantly greater than one, a power function is applicable, and if γ is significantly greater than δ, the exponential weights on the population variables can be inferred to be larger than one.

3. CONCLUSIONS

This paper shows that the so-called "gravity law" of spatial interaction can be logically derived from the economic principle of utility maximization, rather than from the vague and irrelevant concepts of social physics. The number of trips taken from a given origin to a particular destination per unit time is the sum of the numbers of trips per unit time that maximizes the utilities of spatial interaction of the individuals of the origin, subject to some relevant constraints. Previously used spatial interaction models can be improved through more explicit specification and the incorporation of changes in the independent variables over time.

REFERENCES

[1] Anderson, T. R. "Intermetropolitan Migration: A Comparison of the Hypotheses of Zipf and Stouffer," *American Sociological Review*, 20 (1955), 285–291.

[2] ———. "Potential Models and Spatial Distribution of Population," *Papers, Regional Science Association*, 2 (1956), 175–182.

[3] Bechdolt, B. V., Jr. "General Directional Spatial Interaction Models," unpublished thesis, University of Southern California, August, 1967.

[4] Bossard, J. H. S. "Residential Propinquity as a Factor in Marriage Selection," *American Journal of Sociology*, 38 (1932), 219–244.

[5] Carey, H. C. *Principles of Social Science*. Philadelphia: Lippincott, 1858–1859.

[6] Carroll, J. D. "Spatial Interaction and the Urban-Metropolitan Description," *Papers, Regional Science Association*, 1 (1955), D1–D14.

[7] Carrothers, G. A. P. "An Historical Review of the Gravity and Potential Concepts of Human Interaction," *Journal of the American Institute of Planners*, 22 (1956), 94–102.

[8] Clark, C. and G. H. Peters. "The 'Intervening Opportunities' Method of Traffic Analysis," *Traffic Quarterly*, 19 (1965), 101–119.

[9] Converse, P. D. "New Laws of Retail Gravitation," *Journal of Marketing*, 14 (1949), 379–384.

[10] Dodd, S. C. "The Interactance Hypothesis: A Gravity Model Fitting Physical Masses and Human Groups," *American Sociological Review*, 4 (1950), 245–256.

[11] Dunn, E. S., Jr. *The Location of Agricultural Production*. Gainsville: University of Florida, 1954.

[12] ———. "The Market Potential Concept and the Analysis of Location." *Papers, Regional Science Association*, 2 (1956), 183–194.

[13] Harris, C. D. "The Market as a Factor in the Localization of Industry in the United States," *Annals of the Association of American Geographers*, 44 (1954), 315–348.

[14] Hoover, E. M. *The Location of Economic Activity*. New York: McGraw-Hill Book Co., Inc., 1948.

[15] Huff, D. L. *Determination of Intra-Urban Retail Trade Areas*. Los Angeles: Real Estate Research Program, Graduate School of Business Administration, Division of Research, University of California, 1962.

[16] ———. "Defining and Estimating a Trading Area," *Journal of Marketing*, 28 (1964), 34–38.

[17] Isard, W. *Location and Space Economy*. Cambridge: The M.I.T. Press, 1956.

[18] ———, *et al. Methods of Regional Analysis*. New York: The Technology Press of MIT and John Wiley and Sons, Inc., 1960.

[19] ——— and G. Freutel. "Regional and National Product Projections and Their Interrelationships," *Long Range Economic Projection, Studies in Income and Wealth*, Vol. 16, Conference on Research in Income and Wealth, for National Bureau of Economic Research. Princeton: Princeton University Press, 1954.

[20] Lösch, A. (W. H. Woglom and W. F. Stolper, trans.) *The Economics of Location*. New Haven: Yale University Press, 1954.

[21] Quandt, R. E. "Some Perspectives of Gravity Models," in *Studies in Travel Demand*, Prepared by MATHEMATICA for the Department of Commerce under Contract No. C-247-65 Neg., 1965.

[22] Quandt, R. E. and W. J. Baumol. "The Demand for Abstract Transport Modes: Theory and Measurement," *Journal of Regional Science*, 6 (1966), 13–26.

[23] Ravenstein, E. G. "The Laws of Migration," *Journal of the Royal Statistical Society*, 47 and 52 (1885 and 1889).

[24] Reilly, W. J. *Methods for the Study of Retail Relationships*. University of Texas Bulletin No. 2994, 1929.

[25] Stewart, J. Q. "An Inverse Distance Variation for Certain Social Influences," *Science*, 93 (1941), 89–90.

[26] ———. "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," *Geographical Review*, 33 (1947), 461–485.

[27] ———. "Suggested Principles of Social Physics," *Science*, 106 (1947), 179–180.

[28] ———. "Demographic Gravitation: Evidence and Applications," *Sociometry*, 6 (1948), 31–57.

[29] ———. "Concerning Social Physics," *Scientific American*, 178 (1948), 20–23.

[30] ———. "The Development of Social Physics," *American Journal of Physics*, 18 (1950), 239–253.

[31] ———. "A Basis for Social Physics," *Impact of Science on Society*, 3 (1952), 110–133.

[32] Stouffer, S. A. "Intervening Opportunities: A Theory Relating Mobility and Distance," *American Sociological Review*, 5 (1940), 845–867.

[33] Von Thünen, J. H. *Der Isolierte Staat in Bezienhung auf Landwirtshaft und Nationalö konomie*. Hamburg, first vol. (1826), third vol. and new edition (1863).

[34] Voorhees, A. M. "A General Theory of Traffic Movement," *Proceedings of the Institute of Traffic Engineers*, (1955), 46–50.

[35] Weber, A. *Über den Standort der Industrien*, Part I. Tubingen. (1909). Also, Friedrich, C. J. trans. and ed. *Alfred Weber's Theory of Location of Industries*. Chicago: University of Chicago Press, 1928.

[36] Young, E. C. *The Movement of Farm Population*. Bulletin 426. Ithaca: Cornell Agricultural Experiment Station, 1924.

[37] Zipf, G. K. "The Hypothesis of the 'Minimum Equation' as a Unifying Social Principle: With Attempted Synthesis," *American Sociological Review*, 12 (1947), 646–647.

[38] ———. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge: Addison-Wesley Press, 1949.