URBAN TRANSPORT EXPENDITURES

by Shlomo Angel and Geoffrey Hyman*

I. INTRODUCTION

Expenditures on transport in urban areas have been the major explanatory variables in spatial economic theories of the city. In particular, they have been considered to be the key to the determination of rents, densities, and land uses in the city. The basic observations underlying the role of transport expenditures in the determination of rents are that central locations in cities require low expenditures on transport and command high rents, while peripheral locations require high expenditures on transport and command lower rents. In addition, densities are found to be high in areas of low transport expenditure and high rents, and low in areas of high transport expenditure and low rents. These observations have existed in the economic literature for many years. Recent work in the field has been more concerned with attempting to integrate these observations into a consistent theory, with the development of mathematical models of the urban land market. Most prominent in this respect have been the works of Alonso [1], Mills [8] [9], Muth [10] [11], and Wingo [14]. All of these models attempt to explain the spatial structure of the city from economic assumptions coupled with some representation of urban transport expenditures. The economic assumptions made by these authors differ in several important aspects, and will not be reviewed here. In this paper we wish to focus on their treatment of urban transport.

Clearly, all urban economists lay great stress on the Central Business District (CBD). The tradition followed by these economists is to assume that all employment is concentrated at a central location, which is usually represented as a point. All travel is then assumed to terminate at this single employment center. This tradition owes its origin to the theory of agricultural production proposed by von Thünen [12] in his "Isolated State," first published in 1826. In the isolated state, all goods are sold at a single market center, distance from that market being a major determinant of land rent. The assumption of a single employment center has been maintained by the urban economists in the face of increasing decentralization, and a decline of the role of the central business district as the single focus of productive activity. It has survived despite its increasing irrelevance because it produces simplifications in the analysis and permits important results to be derived which would be difficult to obtain otherwise. It will be shown here that this assumption leads to unrealistic conclusions about the spatial structure of cities. Ultimately, it must be rejected in order to remove inconsistencies from the theory. The relax-

* The authors are associated with the Centre for Environmental Studies, London, England.

ation of the assumption of a single employment center poses several problems. When employment is distributed throughout the city the expenditure on travel to work from a given residence is dependent on the choice of workplace. Our problem is to determine a transport expenditure function, given a distribution of workplaces and a representation of the transportation system. This transport expenditure function must represent an average expenditure on travel to work, the average being taken over the distribution of workplaces corresponding to a given place of residence.

In this paper we derive this transport expenditure function and show that it possesses the properties necessary to overcome the difficulties inherent in the view of transport which has been postulated by the urban economists. It is important to note, however, that while the economists consider money spent on transport as the variable determining the location of the household, we have adopted time expenditures, rather than money expenditures, as the determining variable. Since our analysis is restricted to car travel only, and does not consider public transit, travel time appears to be the best proxy for the variable portion of transport expenditures. All other expenditures associated with transport: fuel, wear and tear, effort and the like are closely related to it, and should therefore be represented with sufficient precision by travel time. Throughout the analysis, therefore, our reference to transport expenditures will not distinguish between money expenditures and time expenditures. In a similar manner, cost and time will be used interchangeably. The paper is divided into two parts. In the first part (Section 2), we present the economists' postulated structure of urban transport expenditures, and compare it (in Section 3) with our derived form of urban transport expenditures. In the second part (Section 4), we discuss how this form was derived. The discussion of the derivation of the transport expenditure function will be condensed. It is based in large part on a model of urban spatial interaction developed in a previous paper; see Angel and Hyman [4].

This model is a continuous analogue of the "maximum entropy" gravity model derived by Wilson [13]. Its function is to allocate commuters among workplaces given information on the distribution of residences and workplaces, and the spatial separation between them. We present density functions for residences and workplaces of commuters and a velocity field, all calibrated for Greater Manchester in 1965. The velocity field represents the transportation system of the city. It provides information on the cost of moving a unit distance at any location in the city, taking into account the variations in traffic conditions at different locations. Then, when travel is assumed to take place on minimum time paths, we can obtain measures of the travel time between any pair of points. The subject of travel on minimum time paths has been discussed in detail by Angel and Hyman [3]. The transport expenditure function is then derived from the model as the average expenditure on commuting to work from a given location in the city, given the fact that residents at that location distribute themselves among many workplaces. It is important to point out here that, while in the traditional economic analysis the transport expenditure function is usually assumed to be given, and the distribution of residences is to be derived, the situation in our case is reversed. The distribution of residences is assumed to be given, and the transport expenditure function is derived as an average expenditure of residents in a given area. This is a significant difference which cannot be overlooked. While in the economists' theories, transport expenditures determine the overall location of residences, we consider the overall location of residences to determine transport expenditures. Clearly, there need be no inconsistencies between these viewpoints and the integration of the two theories awaits further developments.

The spatial framework of the analysis remains basically the same as that of the urban economists, with one important difference. We have preserved the view of the city as a continuous terrain with all quantities under discussion regarded as functions of distance from a single center. To this extent we follow their tradition. This assumption appears justified by several observations which indicate the importance of distance from the city center in the determination of urban spatial structure. Mills [9; p. 238], for example, quotes a study by Brigham [5] which compared three accessibility measures by correlating them with land values in Los Angeles. The three measures were airline distance to the CBD, roadway distance to the nearest freeway exit, and some measure of employment potential (employment levels weighted by the inverse of the distances to employment centers). Brigham found these measures to be highly correlated with each other and that airline distance to the CBD explained most of the variation in land values. When distance from the center plays such an important role, even in a city as decentralized as Los Angeles, we feel justified in retaining it as a major explanatory variable. The assumption of radial symmetry and the display of all information as a function of distance from the city center will be maintained throughout the analysis. However, we do not wish to retain the assumption that all employment is concentrated in the CBD. Thus, while the subject of the economists' discussion reduces to a single line, our discussion must encompass the two-dimensional plane of the city. Travel can now take place between any pair of points. With this basic difference in mind, we turn to the urban economists' view of transport expenditures.

2. A CRITIQUE OF THE URBAN ECONOMISTS' VIEW OF TRANSPORT

Alonso, Mills, Muth, and Wingo are concerned with equilibrium in the city's land market. They all devote a major part of their work to the spatial structure of the housing market. In deriving the optimum location of households, they consider the substitutability of rent and transport. Households are willing to pay more for rent when they have to pay less for transport, and less when they have to pay more for transport. Alonso, Muth, and Mills derive the equilibrium location of households by equating marginal expenditures on rent and transport. Mills deduces the form of the urban rent gradient by equating the value of the marginal productivity of land used for transport with land rent. Clearly, in all these models the form of the transport expenditure function plays a crucial role. The different authors make different statements about it, although it can be shown that they all share very similar views. Given that all these authors consider workplaces to be concentrated at a single point in the city center,[1] we can infer that commuting costs at the center are zero, since people living there do not have to spend anything on travel to work. Furthermore, since all destinations are concentrated there, it is only natural to assume that commuting costs increase with distance from the city center. The latter assumption is stated clearly in all the works mentioned above.

Mills, Muth, and Wingo further assume that commuting costs increase at a decreasing rate with distance from the city center. This assumption is based on the common observation that congestion in central areas reduces speeds and increases costs per unit distance. To quote Mills [8; p. 204], "travel is inevitably slower in denser, higher rent areas, even in an optimum transportation system." A similar statement is made by Muth [11; pp. 19–20]. "If anything, traffic generally moves more rapidly at greater distances from the CBD; less time in transit is thus spent per mile and vehicles, up to a point, operate at more economical speeds with less stopping and starting. Therefore, I assume that the variable portion of transport costs increases at a non-increasing rate with distance from the CBD or any other center of activity." Wingo [14; pp. 96–97] is more explicit and provides a more strict form for the transport expenditure function. "We can summarize the technical conditions by describing the common transportation function as in Figure 26 [reproduced in this paper in Figure 1]. In this function FH we have assumed a constant operating velocity represented by the constant slope FH, and an average ingression loss, all of which occurs at the center, represented by OF. An alternate assumption would distribute ingression losses throughout the system, yielding a function such as OGH, which implies a decline in average velocity as one moves from G into increasing congestion until the center is reached. It is the OGH type of function which will be assumed hereafter." Wingo [14; p. 101] later remarks, "The curve OGH is probably more representative of the transportation function as generally experienced in urban areas."

It can be easily seen that the general form of the transport expenditure function described by Wingo is implied in the analysis of the other economists. Since all workplaces have been assumed to be concentrated at the city center, it must have a value of zero at the center. The people living and working at the center should not be spending anything on transport. It increases at a non-increasing rate, since its gradient is the cost per unit distance, which is assumed to decrease with distance from the center. It must also be asymptotic to a straight line with a

[1] Mills attempts to avoid concentrating all workplaces in the city center and distributes some workplaces throughout the city. However, oddly enough, these workplaces where output is produced do not require commuter transportation. Mills [9; p. 243] states that "the total demand for transportation at u (distance from the city center) is proportional to the output produced beyond u. When output is goods, the assumption can be interpreted to mean that a certain fraction of goods is transported to the city center. When output is housing services, the assumption can be interpreted to mean that a certain amount of commuter traffic is generated per household." Mills appears to ignore the demand for transportation by commuters who produce output outside the city center. He clearly assumes that all destinations are located at the city center, although output is produced throughout the city. From a transportation point of view, then, this assumption is identical with that of Alonso and others.

positive slope. This follows since velocities cannot increase indefinitely with current transport technology and are therefore bounded from above. After a certain distance from the city center, velocities must level off. Cost per unit distance must therefore level off; hence the constant gradient. The gradient must be strictly positive because the cost per unit distance cannot level off at zero. There is therefore always an additional cost to be accounted for when distance from the center increases.

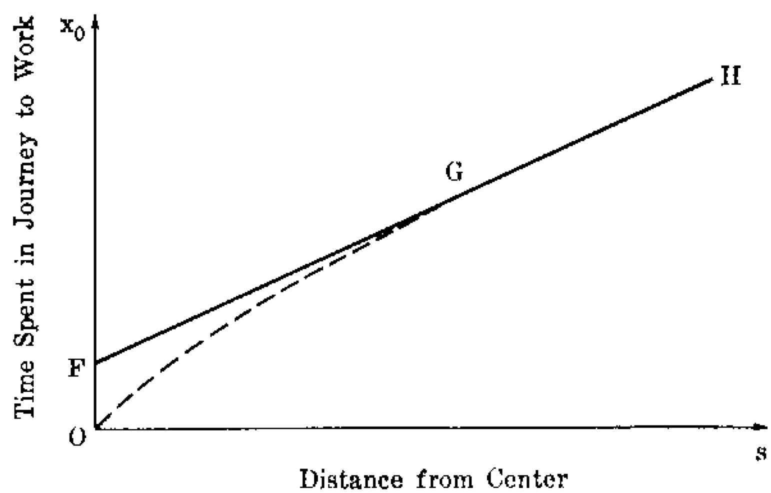

FIGURE 1. Wingo's Transport Expenditure Function

Wingo's form of the transport expenditure function can be taken as the form implicit in the work of the other economists, being derived directly from their assumptions. This form of the function is consistent with our observations using Manchester data for car commuters in 1965, in the case where everyone is assumed to be travelling on radial routes to the city center. We have constructed a radially symmetric velocity field for Greater Manchester, using car travel data for 1965; see Angel and Hyman [2; pp. 212–213]. The velocity of travel at any given distance from the city center was obtained by averaging the velocities on a sample of road links at that distance. A continuous curve was fitted to these values. This curve is presented in Figure 2. It takes the form

$$V(r) \quad 24.9 - 16.9e^{-0.58r} \tag{1}$$

where r is measured in miles and V in miles per hour. We compared travel times on minimum routes in this velocity field (see discussion of minimum routes in Section 4) with travel times on the road network using data from the SELNEC (South East Lancashire-North East Chesire) Transportation Study;[3] see Angel and Hyman [3]. We found these two sets of measurements to be highly correlated and

[3] We thank A. Hawkins and M. Hammerstone of the Mathematical Advisory Unit of the Ministry of Transport for their continuous help in obtaining and interpreting the data.

lying on a straight line with a slope of 0.74. We therefore multiplied the original velocity field by a factor of 0.74 to obtain a new velocity field:

$$V(r) = 18.5 - 12.5e^{-0.60r} . \tag{2}$$

The reciprocal of the velocity at any location, $1/V(r)$, is the time it takes to travel one mile at a distance r from the center. It is convenient to describe the information

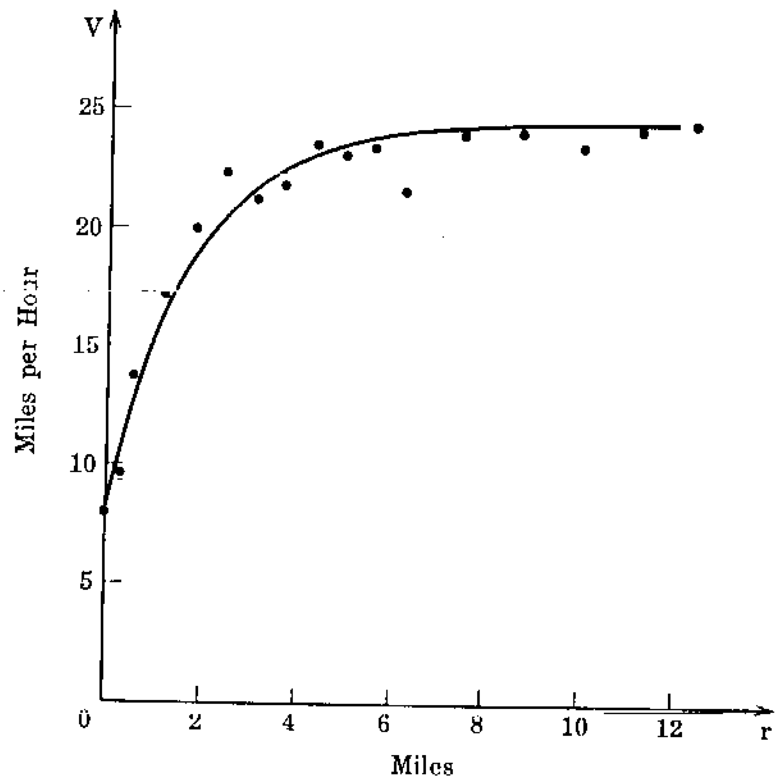

FIGURE 2. Average Velocity as a Function of Distance from Central Manchester, 1965

specified in the velocity field by means of the travel cost field, which measures the time per unit distance at any location. The total travel time along a path is the sum of times spent traversing small distances in the locations through which the commuter travels. For radial travel, therefore, the expenditures on travel to the city center are obtained by integrating the travel cost field between the center and any given location. Expenditures on transport to the center of Manchester are presented in Figure 3. As can be seen by comparing Figures 2 and 3, both curves share the same properties. Both are zero at the origin; both increase at a decreasing rate; and both are asymptotic to a straight line with a positive gradient.

We now proceed to analyze the difficulties inherent in this form of the transport expenditure function, and hence also in the assumption that all destinations are concentrated in the CBD. Wingo [14; p. 65] makes the critical assumption that the sum of expenditures on rent and transport is fixed for any given household. Muth [11; p. 21] and Alonso [1; p. 25] deal with the total household budget, rather than with a special budget for housing and transport. They postulate a composite

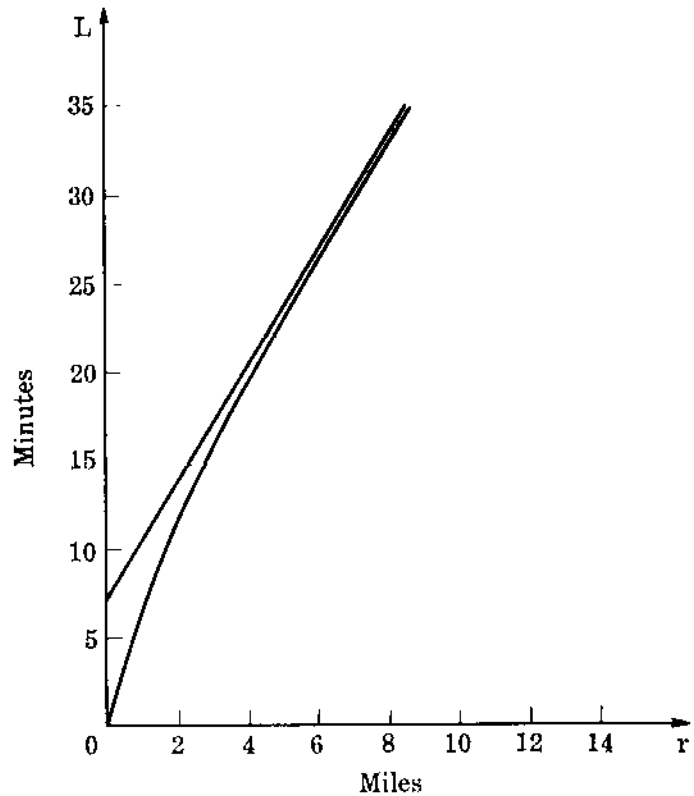

FIGURE 3. Expenditures on Commuting to the Center of Manchester as a Function of Distance of a Given Residence from the Center, 1965

good which includes all goods purchased except housing (or land in the case of Alonso) and transport and which has some given composite price. The household in this case is assumed to make its location decision by maximizing a utility function which depends on the quantities of housing (or land) and other goods purchased, subject to a household budget constraint. In mathematical notation, we maximize $U(q, z)$ subject to the budget constraint

$$y = p_z z + p(r)q + L(r) \tag{3}$$

where:

$p(r)$ = price of a unit of housing (or land) at a distance r from the city center;
q = quantity of housing (or land) purchased;

$L(r)$ expenditure on travel to the center for a household locating at r;
z = is the quantity of the composite good;
p_z = price of a unit of the composite good.

A necessary condition for household equilibrium is then

$$\frac{dL}{dr} \qquad q\frac{dp}{dr} \tag{4}$$

where dp/dr and dL/dr are the rates of change of house prices and transport expenditures with respect to distance from the city center. This condition, which appears in Muth [9; p. 22], Alonso [1; p. 24, Figure 4], and Mills [8; p. 205], requires that the household cannot increase its real income by a change in location. It states that, for a given quantity of housing, the rate of change of housing expenditure is of the same magnitude as the rate of change of transport expenditure, but with the opposite sign. Muth then argues that since transport expenditure increases with distance from the center, house prices must decline with distance. According to Figures 2 and 3, the transport expenditure gradient approaches a strictly positive value at large distances from the center. If we now assume that the quantity of housing purchased is bounded above, i.e., that people cannot buy infinitely large houses, it follows that house prices must decline at a rate bounded away from zero. In other words, it would appear that the house price or the rent function should approach a straight line of strictly negative slope. From these conditions, we can deduce that house prices or rents should become negative at some distance from the center. This conclusion is deduced directly from the first-order condition for equilibrium stated in equation (4), and therefore also applies to the analyses of Alonso and Mills.[3] In the case of the Wingo analysis, where the sum of expenditures on rent and transport is assumed to be fixed, the conclusion is similar. Since transport expenditures increase indefinitely, rents are bound to become negative at some point.

The occurrence of negative rents is usually obscured by introducing a "city limit" into the economic analysis. Urban economists then restrict their attention to the region within this limit, leaving the rest of the country to students of another field. However, this now raises the problem of specifying the appropriate boundary conditions. In particular, we should not expect the house price gradient to change abruptly at some arbitrary distance from the center, nor is there data to support such a contention. It is more reasonable to expect house prices to approach a constant positive value. This, however, would not be consistent with the economists'

[3] Muth [11; p. 72] later modifies his analysis and assumes that transport expenditures increase at a constant rate in order to derive a negative exponential house price gradient. This would imply that transport expenditures must vary linearly with distance from the city center. Since all jobs are located at the CBD, according to his analysis, travel cost must be zero at the center. If we assume that the gradient of the transport expenditure function is positive, rents become negative at large distances from the center. It must follow that the gradient of the transport expenditure function is zero and hence that transport expenditures would be zero everywhere. In the light of these criticisms, Muth's derivation of the negative exponential house price function requires some revision.

transport expenditure function, because it would require the latter to approach a positive constant. If the transport expenditure function is to approach a constant at large distances from the city center, we must abandon the requirement that all destinations are concentrated at the CBD. At large distances from the CBD the price of housing or land must be unrelated to the cost of travelling to the city center.

3. THE TRANSPORT EXPENDITURE FUNCTION

It appears that the economic analysis of cities which requires that all workplaces be located in the CBD leads to unrealistic conclusions about the structures of rents and house prices. We must therefore reject the assumption that all workplaces are located in the CBD and consider a distribution of jobs over the entire city. With such a distribution, it becomes possible to obtain a more realistic transport expenditure function. In particular, a person residing at a distant location could obtain employment close to his place of residence, and is therefore not required to bear the large transport expenditure involved in commuting to the CBD. When this is the case, transport expenditures no longer increase indefinitely with distance from the center. They flatten out as the distance from the center increases, and reach a maximum value of expenditure on transport. With a distribution of jobs, transport expenditures are not uniquely determined at each distance from the city center. Since people work everywhere, transport expenditures at a given location can only be determined as an average expenditure for people living there and working in various destinations throughout the city. When these expenditures are averaged over many workplaces, they are no longer zero at the city center. Expenditures on travel to work from the center are obtained by averaging expenditures for people living in the center and working throughout the city, and can no longer be taken as negligible. Furthermore, when the analysis is restricted to car travel, to the exclusion of public transport, central locations are not necessarily locations of least transport expenditures, as in the preceding economic analysis. Given that people living in the center must travel through congested areas on their way to work, it is no longer necessarily true that central locations are most accessible to jobs.

All of these properties are exhibited in the transport expenditure function derived for Manchester in 1965. This function was derived, as stated earlier, by taking as given the distribution of residences and workplaces of car commuters in Greater Manchester, and by allocating commuters between residences and workplaces in accordance with a continuous trip-distribution model. This model allocates commuters living at a given location to workplaces in proportion to the density of residents in the given location, the density of jobs at the workplace, the relative accessibilities of these locations to jobs and residences, respectively, and some measure of the cost of overcoming the distance separating them. We then obtain the transport expenditure function by averaging the transport expenditures of commuters from a given place of residence. Figure 4 illustrates the transport expenditure function for Greater Manchester, 1965. The vertical axis represents commuting costs, measured in our case in minutes. The horizontal axis represents distance from the city center.

As can be seen from Figure 4, proceeding outwards from the city center, transport expenditures decrease for the first half mile from about 9.5 minutes to about 8.8 minutes. This effect is associated with congestion in the central area, since it

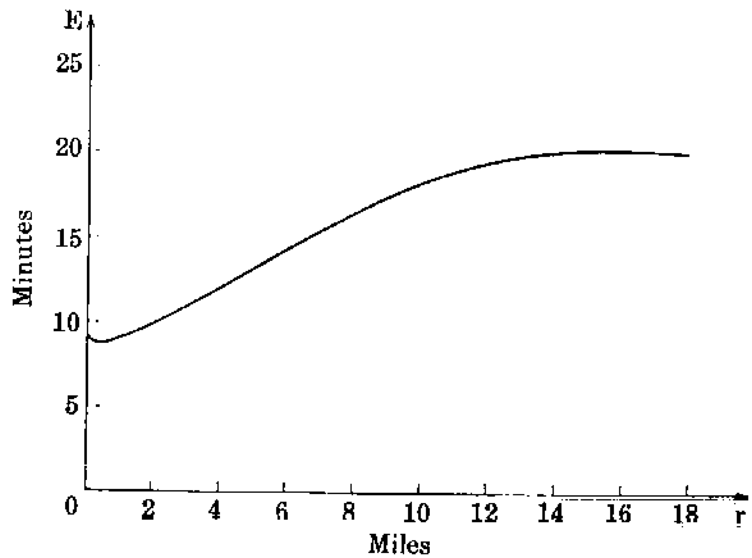

FIGURE 4. Transport Expenditures as a Function of Distance of Residence from Central Manchester, 1965

did not appear when average transport expenditures were derived for a uniform travel cost field. Transport expenditures then increase, at an increasing rate, until they reach a point of inflection at about 4.5 miles out. They subsequently increase at a decreasing rate to reach a value of about 20 minutes at a distance of 14 miles from the center. Transport expenditures then level out and remain at this value for the rest of the range. The latter effect is probably due to the predominance of local jobs at large distances from the center. It is also important to note that transport expenditures in general are far lower than those presented earlier (Figure 3) when only radial travel to the city center is considered. This can be easily seen by comparing Figures 3 and 4, which are drawn at the same scale. The latter is markedly shallower. It must follow, therefore, that the assumption that commuters all travel to the CBD overestimates their expenditures on transport. This can be best illustrated by comparing transport expenditures for people who actually work in the CBD with those who work elsewhere. Such a comparison is shown in Figure 5 which exhibits the expenditures on transport for people working at a given distance from the city center.

As can be seen from Figure 5, within most of the urban area, those car commuters who work in the city center spend most on travel to work. Expenditures on travel to work decline steadily for more suburban workplaces and reach a minimum at a distance of approximately seven miles from the center. Employees working at these distances spend least, on average, on commuting. Jobs located further out require higher expenditures on transport, and the function did not seem to level out within the range of study. Clearly, then, the economists' assumption

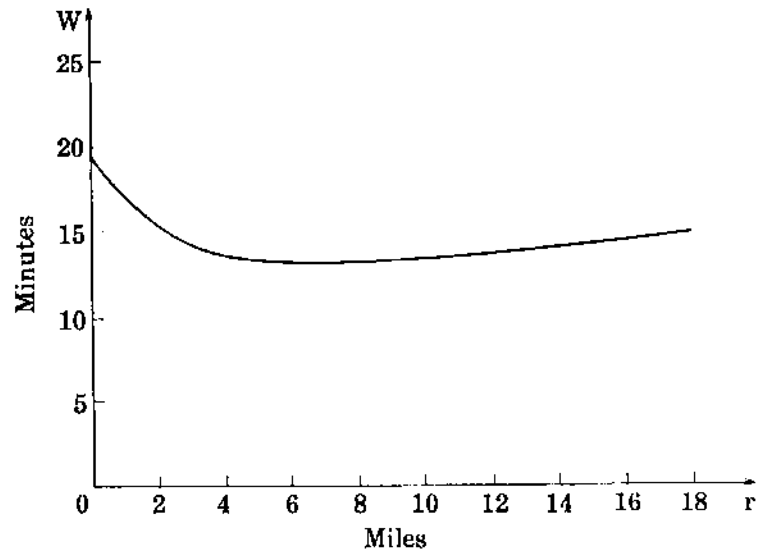

FIGURE 5. Transport Expenditures as a Function of Distance of Workplace from Central Manchester, 1965

that jobs are concentrated at the center considerably overestimates the amount of time or money spent on commuting. Most people will have smaller expenditures on commuting than those postulated by the economists. We must therefore conclude that the economists will inevitably attribute more to the effects of transport expenditures on rent and location than is, in fact, attributable to them when a distribution of jobs throughout the urban region is taken into account.[4]

We return to the transport expenditure function exhibited in Figure 4. The function possesses the correct properties at large distances from the center, as it reaches a maximum and does not go on increasing indefinitely. Near the center, within the first half-mile range, it does not agree with the economists' assumptions because it is decreasing rather than increasing there. If we assume that rents are at a maximum in the city center, due to competition from commercial and other uses which require more central locations, then the necessary condition for household

[4] The high cost of commuting to the center gives us an indication of the forces of decentralization at work in modern cities. From the point of view of the labor force, savings accrue when jobs can be found away from the city center. We must therefore expect the number of vacancies in central areas to be higher and the wages offered to be higher than those in outlying locations. At large distances from the center, workplaces again require large expenditures on commuting. However, in this case, employees must come from distant locations because of the low density of the population.

equilibrium, equation (4), is not satisfied in this range. In this range, both rents and transport expenditures have negative gradients. We must conclude, then, that no household would locate within this range. Remembering that we have restricted our analysis to car users, we can rephrase this prediction by saying that no resident living within this range would commute to work by car. This prediction is clearly borne out by our observations in Manchester. There are no car commuters within this distance from the center. This is illustrated later by the gamma distribution fitted for residences of car commuters (Figure 5) which has the property that the density of residents at the center is zero. At large distances from the center, if this equilibrium condition is to be satisfied, we should expect the rent gradient to level out when the transport gradient levels out. We have no *a priori* reason, however, to expect it to level off at zero, as in the negative exponential function derived in Muth [11] and Mills [9].

The derived transport expenditure function thus differs from the function postulated by the urban economists in several important ways. It is positive, rather than zero, at the origin. It decreases for a short range near the origin, rather than increasing everywhere. It generally increases at a slower rate. Its slope first increases then decreases over the range, rather than being non-increasing everywhere. And finally, the function approaches a constant value rather than increasing indefinitely. All of these properties can be justified on quite simple grounds and are consistent with the observations. It appears, therefore, that future analysis would have to consider a more appropriate form of the transport expenditure function than that postulated and used by the urban economists. In the following section, we elaborate on the derivation of the function presented above from empirical data on densities and traffic conditions, using a continuous model for the distribution of trips among residences and workplaces.

4. THE DERIVATION OF THE FUNCTION FROM A MODEL OF SPATIAL INTERACTION

A transport expenditure function which reflects a more realistic picture of modern urban areas needs to be derived as an average of the amount of time that commuters spend on travel from a given place of residence. Since these commuters work everywhere in the city we need some procedure which will distribute them among the different workplaces in a realistic manner. Roughly speaking, this procedure must utilize information on the availability of jobs and their relative accessibility to the commuter's residence. One such procedure, the trip distribution model, has been used extensively in modern transportation studies. The role of the trip distribution model is to predict the number of trips between the different locations in the city. What is required is an estimate of the number of commuter residences in each area, the number of jobs in each area, and a measure of the spatial separation between each pair of areas, in our case the travel time. The main difference between this type of analysis and the economic analysis discussed above is that here the distribution of residences is assumed to be given in advance and is not derived as a result. It is required in order to derive a more realistic distribution of trips, taking into account that commuters compete with each other for jobs. It is possible, however, to derive a simpler distribution model where only the distribution of workplaces is specified. In this case, we would have to employ some measure of attraction of places. This type of model has been employed by Lowry [7] to distribute the urban population among residences. In the foregoing analysis, we assume that the distribution of residences has been determined. We follow the approach of Wilson [13] who derived a trip distribution model of the type required by also assuming that the average transport expenditure in the entire urban area had been determined.

Wilson poses the following problem: given the above information, what is the most probable distribution of trips? He then derives a most probable distribution by maximizing the entropy of the distribution subject to the following constraints: trips leaving an area must sum up to the number of origins (residences of commuters) in that area; trips arriving at an area must sum up to the number of destinations (workplaces) there; and the sum of all transport expenditures must equal the total budget for transportation. Wilson has formulated this model in the context of transportation studies which invariably use a set of discrete zones as a framework for analysis. We retain here the view of the city as a continuous terrain with jobs and residences given as distributions over geographic space. It is therefore necessary to derive a continuous analog of this model. Such a model has been developed by Angel and Hyman [4], and a program was written for calibrating the model using Manchester data.[6] We now briefly present this model.

Restricting the analysis to radially symmetric cities, we define the origin density function, $O(r_1)$, as the number of trips (in our case car trips) originating in a unit area about a point r_1 miles away from the city center. Similarly, we define the destination density function, $D(r_2)$, as the number of trips terminating in a unit area r_2 miles from the city center. We now define the trip density function, $T(r_1, \theta_1, r_2, \theta_2)$, as the number of trips from a unit area around the point (r_1, θ_1) to a unit area around the point (r_2, θ_2), using a polar coordinate system (r, θ). The travel time between this pair of points will be denoted by $t(r_1, \theta_1, r_2, \theta_2)$. Given these definitions we can now write the continuous analog of Wilson's maximum entropy model. Maximize the entropy

$$H = -\int_0^{\infty}\int_0^{\infty}\int_0^{2\pi}\int_0^{2\pi} T(r_1, \theta_1, r_2, \theta_2) \quad \ln T(r_1, \theta_1, r_2, \theta_2) r_1 r_2 d\theta_1 d\theta_2 dr_1 dr_2 \tag{5}$$

subject to the following constraints:

The total number of trips from a unit area at (r_1, θ_1) to all other areas must be the density of origins there, that is,

$$\int_0^{\infty}\int_0^{2\pi} T(r_1, \theta_1, r_2, \theta_2) r_2 d\theta_2 dr_2 = O(r_1) \quad \text{for all } r_1, \theta_1 . \tag{6}$$

Similarly, the total number of trips originating in the entire city and terminating in a unit area at (r_2, θ_2) is the density of destinations there, that is,

[6] The program is described in full in Angel and Hyman [4; Appendix 4].

$$\int_0^{\infty}\int_0^{2\pi} T(r_1, \theta_1, r_2, \theta_2) r_1 d\theta_1 dr_1 = D(r_2) \quad \text{for all } r_2, \theta_2 . \tag{7}$$

Finally, the total expenditure on transport in the city is the product of the average expenditure, $\bar{t}$, and the total number of trips, T, in the system, that is,

$$\int_0^{\infty}\int_0^{\infty}\int_0^{2\pi}\int_0^{2\pi} T(r_1, \theta_1, r_2, \theta_2) t(r_1, \theta_1, r_2, \theta_2) r_1 r_2 d\theta_1 d\theta_2 dr_1 dr_2 = T\bar{t} . \tag{8}$$

The solution to this maximization problem can be obtained using calculus of variations methods. It yields a trip density function of the form[6]

$$T(r_1, \theta_1, r_2, \theta_2) = A(r_1)B(r_2)O(r_1)D(r_2)e^{-\mu t(r_1, \theta_1, r_2, \theta_2)} . \tag{9}$$

The constraints (6) and (7) can now be rewritten as

$$\frac{1}{A(r_1)} = \int_0^{\infty}\int_0^{2\pi} B(r_2)D(r_2)e^{-\mu t(r_1, \theta_1, r_2, \theta_2)} r_2 d\theta_2 dr_2 \tag{10}$$

and

$$\frac{1}{B(r_2)} = \int_0^{\infty}\int_0^{2\pi} A(r_1)O(r_1)e^{-\mu t(r_1, \theta_1, r_2, \theta_2)} r_1 d\theta_1 dr_1 . \tag{11}$$

Equations (10) and (11), together with constraint (8), can now be used to obtain the forms for the balancing factors $A(r_1)$ and $B(r_2)$, as functions of distance from the city center, and the value of the parameter μ, using numerical methods for integration and iterative techniques for solving the equations simultaneously.

The trip density function thus obtained is proportional to a measure of separation between the two locations, $e^{-\mu t(r_1, \theta_1, r_2, \theta_2)}$. This means that the number of trips falls off at a negative exponential rate as the cost of travel between the points increases. When μ is large, trips fall off at a rapid rate and people do not commute very far. When μ is small, trips fall off slowly and people travel greater distances. The value of μ is heavily dependent on the average expenditure on transport in the system, $\bar{t}$. When $\bar{t}$ is small, people do not travel far and μ is large. When it is large, people travel further and μ is small. The trip density function is also related to the density of origins and destinations at the end points. When there are more residences at a given area or more workplaces at the other end, we can expect more trips to take place between the two areas. The balancing factors, $A(r_1)$ and $B(r_2)$, are functions of location which are associated with accessibility. The reciprocal of the balancing factor for origins, $1/A(r_1)$, is interpreted as the relative accessibility to job opportunities at a location r_1 miles from the city center. From equation (10) it can be seen to be large when there are many jobs in proximity to the location at r_1 and, since it is proportional to the other balancing factor, $B(r_2)$, it is also related by equation (11) to the number of residences of commuters competing for these jobs in the vicinity of r_1. The term $1/A(r_1)$ will therefore be large where there are many potential jobs, and not many residents competing for them. By a similar argument, the reciprocal of the balancing factor for destinations, $1/B(r_2)$, is a measure of the

[6] The proof of this result appears in Angel and Hyman [4; Appendix I].

relative accessibility to residents of a given workplace at r_2. It will be large where there are many residents and few workplaces.

The solution of the model equations makes it possible to derive the transport expenditures at any location. The transport expenditure, $E(r_1)$, at a distance r_1 from the center, is the average expenditure of the residents in a small area there on transport. It is obtained by summing up the costs of individual trips and dividing this sum by the total number of residents there:

$$E(r_1) = \frac{1}{O(r_1)} \int_0^{\infty}\int_0^{2\pi}\int_0^{2\pi} T(r_1, \theta_1, r_2, \theta_2)t(r_1, \theta_1, r_2, \theta_2)r_2 d\theta_1 d\theta_2 dr_2 . \tag{12}$$

The form of the function $E(r)$, the transport expenditure function, derived for Greater Manchester 1965, was presented in Figure 4. In a similar manner, we can obtain the work-based transport expenditure, $W(r_2)$, as the average expenditure on transport for people working in a given location at a distance r_2 from the city center:

$$W(r_2) = \frac{1}{D(r_2)} \int_0^{\infty}\int_0^{2\pi}\int_0^{2\pi} T(r_1, \theta_1, r_2, \theta_2)t(r_1, \theta_1, r_2, \theta_2)r_1 d\theta_1 d\theta_2 dr_1 . \tag{13}$$

The form of this function, derived for Greater Manchester 1965, was presented in Figure 5. To complete the picture, we must now deal with the data inputs used in deriving these functions. As stated earlier, the trip distribution model utilizes two basic forms of inputs: the distributions of residences and jobs of commuters; and a measure of separation between every pair of points, here taken to be the time of travel, $t(r_1, \theta_1, r_2, \theta_2)$. In this particular model we also require a given average expenditure on transport, $\bar{t}$. Empirical data for these were obtained from the SELNEC (South East Lancashire—North East Chesire) Transportation Study. We have chosen to represent these as radially symmetric functions, as is common practice in the urban geographic literature. However, while traditional analyses often employ the negative exponential form of describing urban densities, see Clark [6], we have found this form to be inappropriate for describing our data adequately. In the case of residential densities, the fit was worst near the city center, where there are hardly any residences. In the case of workplaces, we have found that the density declines more rapidly than the exponential. In both cases, we have obtained a better fit with a gamma distribution, having a density function of the form

$$d(r) = ar^b e^{-cr} \tag{14}$$

where $d(r)$ is the density at a distance r from the center, and a, b, and c are parameters to be estimated from the data. In the case of residences, the parameter b takes a positive value; in the case of workplaces it takes on a negative one. Having chosen the gamma distribution as our model, we have obtained maximum likelihood estimates for the parameters a, b, and c for both densities. The resulting density functions are

$$O(r) = 1164r^{0.982}e^{-0.439r} \tag{15}$$

for origins, and

$$D(r) = 4677r^{0.451}e^{-0.298r} \tag{16}$$

for destinations.

Figure 5 illustrates the density function for origins of commuters for the Greater Manchester area in 1965, in isometric form. Figure 7 illustrates the corresponding density function for workplaces. The cost of travel between any two

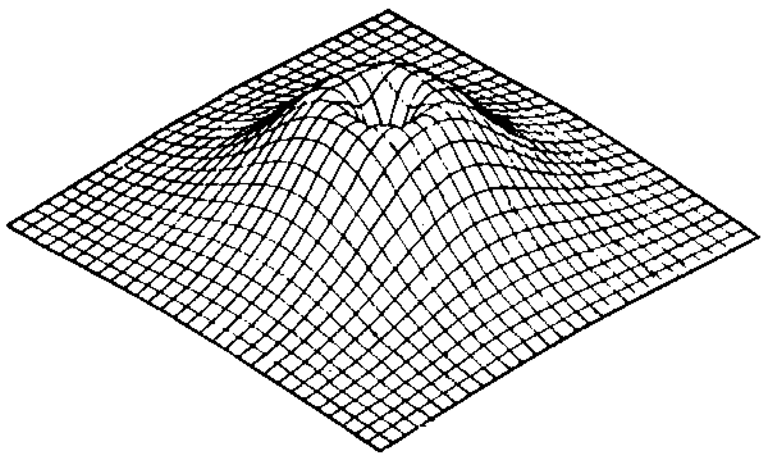

FIGURE 6. Density of Residences of Car Commuters in Greater Manchester, 1965

points was obtained from the velocity field, $V(r)$, for Manchester which was presented earlier. But, while earlier it was only applied to radial travel, we use it in the general case for deriving the expenditure on travel between any two points. We define the cost field as a surface lying above the urban plane such that the height of the surface at each location is the time spent travelling a unit distance there. The two dimensional representation of the cost field for Manchester in 1965 appears in isometric form in Figure 8. If the velocity field were uniform everywhere, the total expenditure on a journey would be proportional to its length, that is

$$E = \frac{1}{V}s \tag{17}$$

where $1/V$ is the cost per unit distance and s the total length of the journey. This equation suggests that we could define a vehicle-mile as a unit of transportation, where $1/V$ specified its price and s the number of units consumed. In commuting, where travel cost is no longer uniform, due to congestion, the marginal expenditure of a commuter at a point on his journey to work is the price he pays for consuming a unit of transportation there, namely,

$$\frac{dE}{ds} = \frac{1}{V}. \tag{18}$$

The latter can be recognized as the differential equation of motion. At any location, the price that the consumer pays will be determined by the interaction between the demand for using the transportation system and the supply of transportation facilities there. The individual is thus faced with a price which is determined by competition. Where many commuters compete for limited road space, the cost per unit

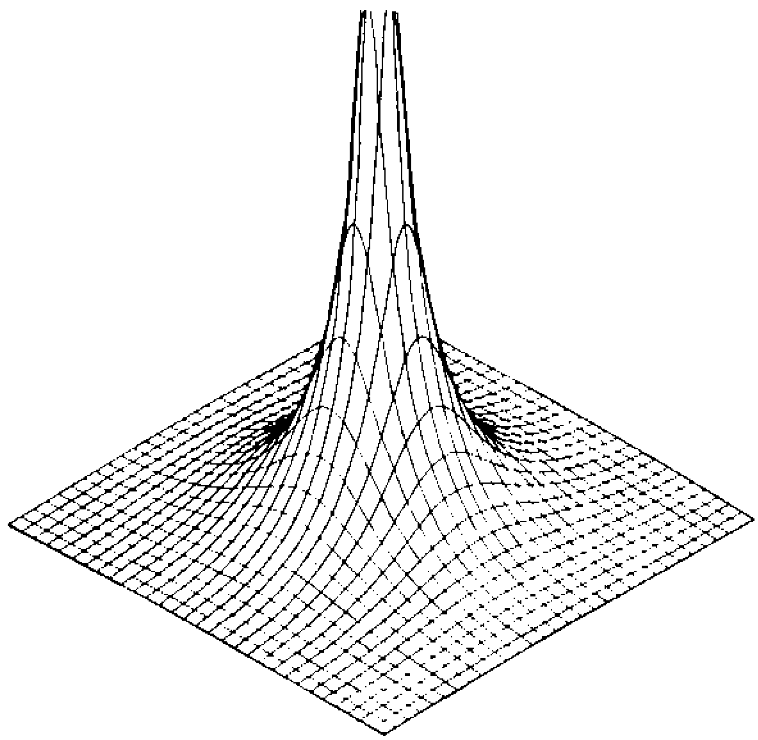

FIGURE 7. Density of Workplaces of Car Commuters in Greater Manchester, 1965

distance is high; where fewer compete, the cost per unit distance is lower. However, each individual commuter travelling through an area cannot significantly affect the cost per unit distance at that location.

The total expenditure on travel along a given path P between two points is given by the path integral under the cost field. It equals the vertical area between the surface and the path P traversed in the urban plane. Any pair of points can be connected with many different paths. There is, however, one path which minimizes the total expenditure on travel between them. If we assume that the commuter will choose this path, we can determine how much he will spend on commuting between these points. This expenditure equals the minimal area between the cost field and the path traversed. A soap film, for example, will naturally form this minimal area when the surface and the plane, connected by two vertical chords at the end points, are dipped into a soap solution. A numerical solution to this problem has been obtained using the calculus of variations, and is described in detail by Angel and Hyman [3]. Finally, the average expenditure on travel in Manchester was given as 14.7 minutes. Given these data, then, it was possible to derive the

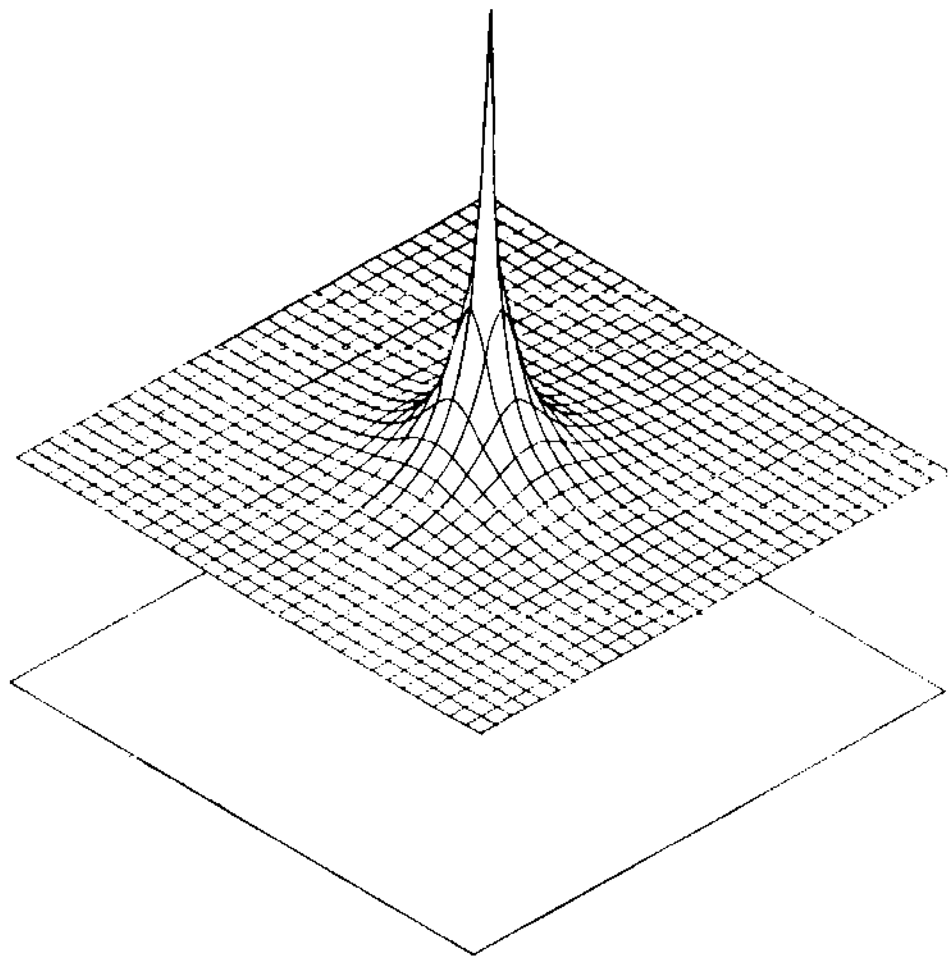

FIGURE 8. Cost Field for Greater Manchester, 1965

form for the transport expenditure function. The effect of changes in the model parameters on the transport expenditure function merits further study and this has not been attempted here. We have satisfied ourselves, therefore, with the discussion of its general properties, rather than with the specific values obtained. It is our hope that this discussion will shed some light on the role of transport expenditures in studies of urban spatial structure.

REFERENCES

[1] Alonso, W. *Location and Land Use.* Cambridge, Mass.: Harvard University Press, 1965.

[2] Angel, S. and G. M. Hyman. "Urban Velocity Fields," *Environment and Planning,* Vol. 2 (No. 2, 1970), pp. 211–224.

[3] ——— and ———. "Urban Travel Time," *Papers of the Regional Science Association,* Vol. 26 (1971), pp. 85–99.

[4] ——— and ———. "Urban Spatial Interaction," *Environment and Planning,* Vol. 2 (No. 4, 1971), pp. 99–118.

[5] Brigham, E. F. "The Determinants of Residential Land Value," *Land Economics,* Vol. 41 (November 1965), pp. 325–334.

[6] Clark, C. "Urban Population Densities," *Journal of the Royal Statistical Society,* Series A, Vol. 114 (1951), pp. 490–496.

[7] Lowry, I. *Model of Metropolis.* Santa Monica: RAND Corporation, RM-4035-RC, 1964.

[8] Mills, E. S. "An Aggregative Model of Resource Allocation in a Metropolitan Area," *American Economic Review,* Vol. 57 (May 1967), pp. 197–210.

[9] ———. "The Value of Urban Land," in H. Perloff (Ed.), *The Quality of the Urban Environment.* Baltimore: The Johns Hopkins Press, 1969.

[10] Muth, R. "The Spatial Structure of the Housing Market," *Papers of the Regional Science Association,* Vol. 7 (1961), pp. 207–220.

[11] ———. *Cities and Housing.* Chicago: University of Chicago Press, 1969.

[12] von Thünen, J. H. *Von Thünen's Isolated State.* Translated by C. M. Waartenburg and edited with an introduction by P. Hall. Oxford: Pergamon Press, 1966.

[13] Wilson, A. G. "A Statistical Theory of Spatial Distribution Models," *Transportation Research,* Vol. 1 (November 1967) pp. 253–269.

[14] Wingo, L. *Transportation and Urban Land.* Washington, D.C.: Resources for the Future, 1961.