Title

Toward a Theory of Urban Public Facility Location

Author
Teitz, Michael B.

Publication Date
1967-11-01

Peer reviewed

WP# 67.

TOWARD A THEORY OF URBAN PUBLIC FACILITY LOCATION

Michael B. Teitz
Department of City and Regional Planning and
Center for Planning and Development Research,
University of California, Berkeley

Working Paper No. 67

Paper prepared for the Fourteenth U.S. Annual Meeting of the Regional Science Association, Cambridge, Mass., November 3-5, 1967.

This research was supported in part by the Economic Development Administration, U.S. Department of Commerce. I would like to thank members of the faculty and students at the Department of City and Regional Planning, University of California, Berkeley, for criticisms, suggestions, and information.

Location theory has come a long way from von Thünen and Weber. In recent years we have witnessed its formalization and codification by Isard [6] promptly followed by two kinds of radical change. The first of these, typified by formal theories of consumer residential choice and advances in analysis of urban retail location, comprises an expansion of the theory's substantive domain. The second change consists of the adoption of computable models, ranging from hill climbing algorithms for the classic Weber problem to complex land use models in transportation studies. Although the latter may be a shift primarily in technology, it appears to be so far reaching that its inevitable feedback to theory has hardly begun to take shape.

Both of these developments contribute to the subject of this paper. Its aim is to suggest a further area of substantive concern for location theory, namely, urban public facility location, the application of which will call for computable models.

While it is not entirely true to say that location theorists have ignored the problem of public facility location in cities, their concern has been slight. Isard [6:182 ff.], for example, briefly mentions urban public facilities in the context of a discussion of urban scale economies, but does not consider their location at all. Elsewhere [7], he deals with them as instruments for participation. Other locational analysts have usually dealt with public facilities by assuming either that their location could be considered determined by residential location, or a special case of commercial and retail location. Even in the modern development of central place theory, consideration of public facilities has been marked by virtually no concern with their public rather than market mechanism for determination of scale and

and location. Berry [2] discusses planning applications of central place theory, but cites only one instance of rationalization of public services, and that in a non-urban and non-theoretical context. Similarly, Berry and Pred [3] give virtually no theoretical references pointing toward public facilities, and, although a section of their bibliography is devoted to medical services, no mention is made of others.

On the other hand, the problem has not been totally unconsidered. The relationship between public welfare concerns and the minimum average cost solution to the Hotelling ice cream vendor problem has often been noted. Tiebout [13], in an important paper, outlined some major components for a theory of urban public expenditure decisions, but assumed away most location problems. Suggestive papers have appeared in the recent literature in local public expenditure economics, especially in Schaller [9]. Haggett [5:248-250] cites empirical work on school location and district boundaries. Alexander [1] has put forward some suggestive ideas on urban highway spacing. And finally, in this brief catalogue, the literature on medical facilities offers a number of recent works, notably by Schneider [10].

1. Why a Theory of Urban Public Facility Location?

The failure of location theorists to consider publicly-determined facilities is quite surprising when one comes to consider the role of those facilities in shaping both the physical form of cities and the quality of life within them. Modern urban man is born in a publicly financed hospital, receives his education in a publicly supported school and university, spends a good part of his life travelling on publicly built transportation facilities, communicates through the post office or the quasi-public telephone system, drinks his public water, disposes of his garbage through the public removal system, reads his public library books, picnics in his public parks, is protected by his public police, fire, and health systems; eventually he dies, again in a hospital, and may even be buried in a public cemetary. Ideological conservatives notwithstanding, his everyday life is inextricably bound up with governmental decisions on these and numerous other local public services.

Implicitly, we seem to have assumed as theorists that the character and location of public facilities simply reflect the overwhelming non-public decisions on residential, commercial, and industrial location. Furthermore, we have assumed that since market-based competitive models of the latter should lead to Pareto optimal location patterns under atomistic decisions, the public decisions accompanying them should likewise be optimal. It is sometimes hard to take market-based location theory seriously in the real world. Even if it does hold, one is hard put to see why the political process should respond in an optimal way. Yet, by and large, government has in the past responded sufficiently well to give rise to virtually no theoretical speculation about the nature of the response and its implications for the city.

Urban planners, for all their shortcomings, have recognized the problem in practice. But the lack of a theoretical basis for action has left their response mechanical and inadequate. Rules of thumb about size and location of facilities have been developed, but for the most part without ways to evaluate the results or to stimulate invention of new systems. With services locked in mazes of standards, and without powerful competitive pressures for innovation, only crisis has precipitated demands for reevaluation of urban services. And even then, the response tends to be a shotgun scatter of proposals--not a remarkable result, given that no framework for evaluation exists. In this context, it is interesting to note the current urban crisis in America. With it has come the usual plethora of proposals, this time including measures emanating from the Federal level that could drastically change the pattern of urban public services. The schools are the first target, but no doubt others will face similar problems.

Beyond these questions lies a further uncharted territory. Although theorists may have assumed that public facilities followed private decisions, a powerful feedback of public decisions on private also seems likely. If government can use public facilities as instruments to shape urban growth and social and economic behavior, then a new level of evaluation is superimposed on the usual considerations for public services. Of course, planners have been advocating this for many years, but only now does it begin to seem a serious possibility for reasonably scientific investigation. However, even if such larger aims are not now achievable, the drive toward PPB, systems analysis, and cost-benefit at the local level suggests a rising interest in more efficient and effective utilization of resources in urban services to achieve the direct public ends for which those services were instituted. For this alone some theoretical structure might be invaluable.

2. Public and Private Location

These considerations suggest that urban public facility location is worth some theoretical exploration. They do not demonstrate that such a theory is unified, feasible, or significantly different from other components of location theory, including central place theory. In the remainder of this paper, it is scarcely possible to develop a full blown theory of public facility location. Rather, we shall try to indicate its unique characteristics and sketch out some of its most interesting component problems.

The first question that we shall examine concerns the relation of a theory of public location to conventional location theory. There are several important qualitative differences between them. Both location and the scale of expenditure on urban public facilities are determined by some sort of public or quasi-public process. This central fact gives rise to some important characteristics. It suggests that public location theory may bear a relationship to conventional location theory similar to that between the expenditure side of public finance or welfare economics and conventional economic theory. In each case, the former focusses upon public decisions and government budgets in response to some welfare criterion in a mixed market-nonmarket setting. The latter emphasizes the role of choice, taste, and utility or profit maximization in a predominantly market context. Under assumptions of competition and many decision units, the result should be Pareto optimal allocations. Public expenditure theory constantly seeks some equivalent framework in a situation of one or few decision units and restricted personal choice.

One should not carry the analogy too far. Several writers on local public expenditures have pointed out the asymmetry between national and local public finance resulting from large numbers of governments at the local level. Tiebout [12] used this property of local finance as a basis for a consumer behavior oriented theory of local public expenditures. We recognize the existence of many local governments, but for purposes of analysis will begin by treating public facility location on the basis of a single government.

If a single government is responsible for provision of some service over a given area, then a second powerful contrast with conventional location theory emerges. Location theory of the firm or consumer asserts that many decentralized decisions have produced any observed pattern of locations. The pattern may be described positivistically as a system by central place theory, but its locational efficiency is not usually evaluated in system terms. For example, retail location analysts may count the number of supermarkets in an area and describe their pattern. Usually they do not ask whether the individual markets are too small or too large. If the market is operating effectively, such questions are automatically taken care of.

For public facilities, in contrast, these are central questions. Even if we accept an incrementalist view of government budgeting, the problems of the appropriate scale and location of new facilities remain. If a more general position is adopted for theoretical investigation, the overall scale of operation, i.e., the determination of the budget, becomes critical. Such a view has powerful implications for locational analysis. The theorist of public facility location finds himself inevitably drawn away from the problem of location of the individual facility and toward the structure and location of the entire system of facilities within the area over which the government exerts jurisdiction.

Location theory has not concerned itself much with multiple location systems. While operations researchers have analyzed multifacility location problems in practice, location theorists have dealt little with the theory of the multiple facility firm. Yet many interesting problems arise as soon as one begins to look at location from a system viewpoint. For example, one might ask for an appropriate retail location strategy for supermarket chains of different sizes. Or, again, what happens to the Hotelling problem under multiple unit assumptions? To my knowledge there is very little theoretical discussion of these or similar questions.

System location is not a single problem, as will be shown below. But, in one form or another, it does appear crucial in most considerations of public facilities. Before we go on to examine in detail some of these manifestations, some further characteristics of public facility systems and decisions need to be identified.

2.1 Public Facility Systems

In order to talk about public facility location it is desirable to characterize public facilities. Presumably, we mean those components of the city whose primary function is to facilitate the provision of goods and services declared to be wholly or partly within the domain of government. The range of such goods and the degree of government involvement vary greatly over time and place. Almost every function has been run at some time or place by government. For purposes of the present discussion, we will take the set of governmental urban functions in Western countries, particularly the U.S. and Canada, as our purview. No attempt will be made here to establish a list of such functions or their associated goods and services in detail. In the understanding of urban systems such a list may eventually be necessary and might bear a relation to public location similar to that of the complete classification of industry to the theory of industrial location. Or perhaps a better analogy might be the relation of anatomy to human biology.

For preliminary discussion, some first approximations to classes of functions and facilities may be invaluable. We will approach this through the description of generalized ways of making such distinctions. A first cut at the problem follows the theory of public expenditures in classifying outputs and associated facility systems as public or collective use goods, zero short-run marginal cost goods, or merit goods. An air pollution monitor system is of the first kind, a non-congested street network of the second, and, so far as one can tell, a public library system of the third. This is a well understood basis for discrimination and may have significant utility in the study of location. But for the moment we shall not explore it further here except to note its influence on the classifications below.

A second distinction applies to facility systems rather than to goods and services and focusses on their geometric properties. Since theorization and model building imply abstraction, we would do well to look at the most likely abstractions of facility systems. They appear to take two common forms, point patterns and networks. Point patterns characterize a variety of distributive services where the final phase of distribution is flexible and intermittent. Medical centers, post offices, libraries, police and fire systems are of this type. On the other hand, many services call for continuous connections in space. For these a network characterization seems more suitable. Examples might be water, sewer, electric power, gas, telephone, and highways. Let it be emphasized that these are merely suggestive distinctions. Clearly, for any urban function the abstraction is technologically influenced and often more properly described as a mixture than a pure form of one or the other. In the present context it is useful because we will be concerned more often with point-representable than network-representable systems.

Related to the formal geometric property is another quality which combines geometry and behavior. It often appears useful to characterize point-representable facility systems by the direction of interaction associated with each point in the system, and by the abstraction of its relationship to the city, itself abstracted as a point pattern or surface. A system of facilities representable by a point pattern may take this particular form as a compromise between scale economies in operation and economies in distribution. However, there are interesting differences observable in the distribution process itself. We might distinguish on the basis of direction of flow, for example, among systems involving physical movement, between those that are distributors and those that are collectors. In a formal sense, though, these abstract to the same thing. Rather more interesting might be a distinction among distributional systems on the basis of the formal representation as mappings of their relationships with the city. If we represent the city as a point set or surface, the relations between the facility system as domain and its sphere of concern as range may be one-one, one-many, or many-many. Which of these is appropriate seems to be a function of the technical properties of the system, its status as a public good, and the role of consumer choice in consumption of its output. An example involving the first and third of these is offered by the post office. The collection and distribution facilities of the post office are asymmetric. Ideally, each letter entering the system must find a unique destination, namely its addressee, and all letters must be delivered. To achieve these ends effectively, and retain security, an obvious solution is to assign a unique one-many mapping between post office substations and addresses. Thus, at the last stage of its trip, any letter to a particular address comes from exactly one substation. Admittedly, other methods could be employed and perhaps should be examined, but this one appears to be universal. In contrast, the collection arm requires no such form. Because all letters must go into a common sorting process, there is no need for any letter to enter at a unique point. Additionally, part of the collection cost may easily and without significant loss of security be passed on to the consumer. The result is a mailbox system as the lowest order subsystem on the collection side. It exhibits a many-many relation to the set of addresses in the city, and consequently is not a mapping in the strict mathematical sense.

At the risk of belaboring this point excessively, we might point out that the form of relation determines whether the city must be divided into regions for any particular function. Post office substation delivery regions exist; mail box collection regions do not. Not uncommonly, regions are established administratively with no such functional basis. Large urban public libraries offer examples of such exhaustive and disjoint region systems. They also may exhibit inconsistency in boundaries and general disregard for their existence. Given the character of the library function, this is not surprising, and in fact it is probably more healthy than the creation of such regions in the first place. A similar problem appears in efforts to delineate hospital market areas.

A final characteristic of public facility systems that we must mention is their hierarchical property. Almost universally, hierarchical elements appear in the technical structure of public facility systems. At least two sources contribute to the generation of hierarchies. For spatially continuous or regularly periodic distribution and collection systems, the logic of aggregation and system control calls for attenuation of capacity as the consumer is approached. Thus, electrical distribution requires a decline in voltage which is best attained in discrete steps with a transformer station at each stage. Simultaneously, the number of branches in the network increases. Water and sewer pipe capacities likewise propagate steplike attenuation and dendritic properties.

A similar effect in point-representable systems arises from variation in the functions they perform and concurrent scale variations in requirements necessary to support specific functions, either due to cost or infrequency of demand. This effect parallels the orders of stores and goods in a central place system. Every firehouse may have an engine and ladder company, but a specialized heavy rescue company will be much rarer, although located at some firehouse in the system. Whether the outcome for any system is a pure hierarchy depends on the degree to which the higher order system elements include all functions of lower order. For public services, this may not necessarily be the case, especially where specialized components reflect local peculiarities in the demand structure, such as an alcoholic treatment unit in the local medical center catering to skid row.

Empirical information on hierarchical structure in public facilities is scarce and sometimes misleading. There seems to be confusion between technical requirements and standards. In many cases, the existence of an apparent hierarchy reflects the earlier acceptance of some system of standards that may be of dubious value. Where such standards no longer reflect technology, cost, and social objectives effectively, the observed hierarchy may be quite misleading. Nevertheless, incorporation of hierarchical properties into public location models remains important.

2.2 Public Facility Decisions

Its concern with systems sets off public from conventional location theory in terms of units. The fact of being governmental similarly shifts the emphasis in the structure of optimization, that is, objective functions and constraints.

We have previously mentioned the general problems raised by governmental allocation--the need for a social welfare function, the absence of a competitive price system, and the problem of appropriate levels for allocation and optimization. The response of public expenditure economics to these problems has been cost benefit analysis. In military systems analysis, the heavier emphasis on system design and the impossibility of attaching dollar values to some components of the cost-benefit expression, even though they are quantifiable, has led to a form of cost-benefit in which multiple effectiveness criteria are substituted for the single benefit measure. If we lump these together, calling them cost-utility analysis, they provide a technical framework for decision comparable, if theoretically inferior, to simple profit or utility maximization under competition.

Having said this much, most of the interesting problems still remain. We shall briefly comment on some broad questions first, then pursue those especially germane to public facility location. Governments and their component agencies operate under a budget allocation system. This implies at least three major levels of decision, although they may not be separable in practice. Society decides as a whole what part of its resources will be allocated to objectives whose machinery for realization is administered or heavily influenced by government. Within this overall allocation, government allocates resources to what we will call major program areas. These we will take to be clusters of objectives of a similar type. Within these program areas, in turn, resources must be allocated to particular programs and their associated systems of facilities. At this point, we will not consider the validity of this description in terms of behavior, nor will we speculate the direction of determination, that is, whether the real world budget represents a sectoring of a pie or a plate of cookies. We assume that some allocation process exists whereby a quantity of resources becomes available for a program, loosely defined. This assumption is dangerous in that it tends to lead toward the specification of the system location problem in a constrained budget-effectiveness maximization form. Such a form predicates that the entire budget for the program must be used up, a view consistent with agency behavior, but lacking elegance and clearly inefficient unless exactly the right amount has been allocated to ensure equality of marginal net social return in all programs. The latter, to be known, implies that the pattern of expenditures and benefits is known in each program. Thus the problem is circular. Optimal overall allocation presumes optimal program allocation, which in turn requires an overall allocation for its calculation, especially if interprogram spillovers are taken into account.

We could look at the process as a time-free iteration or a sequential response over budget cycles. In any event, it makes little difference where we enter except insofar as emphasis is placed on achieving optimality at the higher, government-wide, or lower, program-facility system, level. And, of course, it is the latter in which we are chiefly interested. For the former, it might be that a constrained effectiveness-budget minimization form proves more appropriate. On the other hand, the use of multiple non-equivalent effectiveness measures would lead to choice of this form for the system problem.

Another general problem which faces facility location in common with local expenditure allocation is the absence of a social welfare function. Even lacking such a function, we might be willing, if all benefits and costs could be quantified in dollar terms, to accept an aggregate net benefit outcome that appeared large enough. Without monetarily quantifiable benefits, we are inevitably forced to consider in detail the distributional consequences of an outcome both over space and over population groups. Where government policies specify "target groups" this is a desirable result. Yet at the local level it may lead to political complexities beyond manageability.

Given these difficulties, we may now consider relevant variables for public facility system decisions. Evidently, we need to know what it is about systems that influences their cost and effectiveness, however measured. In part, the answer depends on the form in which the public location problem is posed, a topic taken up in section 3 below. In general, however, we may immediately identify the opposing forces of economies of scale and advantages of dispersion.

This is a refrain familiar to location theorists. Of particular interest here is its formulation in system terms in a context which lends itself to effective measurement. For a point-representable facility system, we might hypothesize that the effectiveness of an individual component depends upon its scale and the disposition of the rest of the components in the system. The effectiveness of the system as a whole depends upon the scale of its components and their combined relationship to each other. Under an unconstrained budget, system effectiveness should increase or at least not decrease with the number of components, since new additions do not interfere with the scale of the old. The improvement might increase almost indefinitely, although with decreasing marginal gain. For example, if response time were a measure of fire system effectiveness, then any new fire house would decrease first response time up to the point where every structure in the city became a fire house. If the budget for the system is constrained, this is no longer true.

Under a constrained budget, several variables begin to interact. Let us assume that a distinction is made between capital and operating outlays but that the system structure is highly flexible. The first result of a budget constraint is that addition of a new component to a given facility system reduces the resources available to the previously existing components. Where the resources come from remains indeterminate. We could reduce the capital outlays, i.e., the scale of all or some facilities, or reduce their operating expenditures, i.e., output. However, we must also take into account the interaction between these in turn.

Public facilities produce a variety of goods and services and their relationship to the consuming population varies greatly. Two important sources of variation are individual choice and pricing in the consumption of the outputs. At one extreme are those goods for which consumer choice is virtually non-existent and direct price zero. These include most police and fire services as well as pure public goods. When a man's house is on fire he rarely pauses to consider whether or not to call the fire brigade. Nor, in our present system, is he usually presented with a bill for fire protection service that varies with the size of fire that he enjoyed during the previous year. Thus, once some communal decisions about form and quantity of expenditure on fire service have been made, the form of system will depend upon the expected scale and pattern of fires, the technology of surveillance and response, and the measures of effectiveness employed. If a measure such as response time dominates, then for a given surveillance system the result should be maximum dispersion. However, since there are some widely accepted minimum operating scales--the engine company--dispersion must be tempered by the need to retain this minimum scale of operation. This represents the present pattern of local response in fire systems. Overlaid upon this level are higher order fire suppression and rescue functions that will increase the scale of investment and operations for some system components. Additionally, the fire system incorporates many phenomena that derive from its peculiar relationship to standards established by the underwriters. Since compliance with the standards rather than empirical fire experience determines insurance rates in most cases, their influence on the configuration of facilities is immense. Other criteria also influence fire control facility location, particularly high fire risk and life risk situations such as oil refineries and hospitals, respectively. Thus, for this type of service, the existing pattern of facilities represents a series of compromises plus a certain amount of political and organizational manoeuvering--the more fire houses, the higher the prestige of the chief. Consumer choice in the usual sense plays little part in determining the shape of the system.

Urban services for which consumer choice is significant but price zero present a different picture. Libraries provide a good example of such a function. In general, the larger the scale of a unit in such a system, the more consumers it attracts by virtue of providing better or more varied services, for example by carrying a larger book stock. Since entry is free for members of the community and they commonly travel to the facility for the service, travel cost, including time and inconvenience, represents the major real cost to the consumer. Under these circumstances we would expect to find a distance decay effect on usage. If the budget is fixed but the number of facilities variable, then a larger number of outlets implies a smaller scale at each, but greater aggregate access to the population. The scale effect and the distance effect then conflict. An optimal system must solve for both scale and location.

One further complexity should be noted. We have not defined scale precisely, implicitly assuming it to involve relatively fixed cost components of total cost for any facility. But in public facility systems, as any others, variable costs depend on the volume of services rendered. Since price is zero, this volume depends on demand, which in turn depends on the spatial structure and scale of facilities. The system is in a curious position of being able to generate demand by organizing itself appropriately, but at the same time its variable costs will increase and must be constrained within the budget. Demand and supply are tightly interlocked, indeed so much so that it may be impossible to talk about actual quantity supplied except in terms of the amount demanded.

In practice, many of the influences mentioned earlier temper the form of facility system that emerges. Locations of outlets for "free" services are not easily moved in the face of opposition from citizens who know well that they are not free. The influence of standards pervades system evaluation and perception by participants. Often, little or nothing is known about potential response to changes in location as against other elements that influence demand. A general theoretical form may be extremely useful in pointing up the effect of these distorting factors.

4. Public Facility System Location: Alternative Formulations

The discussion above has purposely avoided precise specification of assumptions and functional forms for public facility location. In this section, we shall specify several versions of the problem with varying degrees of rigor. The objective is both to map out the terrain for further exploration, and to respond to the peculiar demands of a theory dealing with public systems.

A fundamental distinction will be made between static equilibrium and dynamic location problems. Most location theory is of the former type, as one might expect where decision units are small and the question of optimal aggregation has hardly been raised. If the criteria for evaluation are system wide, then dynamic problems make themselves felt quickly. No facility system is built in a day or even a year. The same government that builds and runs today's system expects to be running it next year or ten years hence. Decisions on system structure are bound to be made with an eye to the future, especially if the broader system they serve is expected also to change. Since that broader system is the western city, such an expectation will be the rule rather than the exception. Even so, we shall look carefully at static system problems. They are more familiar; useful methods are available; and they offer the opportunity for finding optimal systems for cities of a given size.

A second distinction will be made between partial and general system theories--in the present context, between the part and the whole. Both individual facility decisions and system wide decisions need to be encompassed for each may be made in practice under conditions of incremental or comprehensive decision processes. Conventional location theory has tended to be partial, notwithstanding Lösch's valiant but unsuccessful attempt to construct a general static spatial equilibrium system for all economic activity. It should be noted, though, that the definition of "general" used here actually encompasses only a single facility system. The most general theory would include all systems simultaneously.

4.1 Static Systems

A first approximation to a static system model draws on Tiebout [13] with considerable modifications. Consider a homogeneous service to be distributed over some spatially distributed given population. We will begin with some formidable assumptions, a standard location theory procedure.

Assumptions:

(1) The service is distributed from a point-representable system of N facilities, each having an identical scale S measured in terms of

capital outlay or some other measure.

(2) The service is consumed by individuals who travel to the facilities for this purpose.

(3) The service is zero priced.

(4) A single system-wide measure of effectiveness, namely, total consumption Q of the service, is specified.

(5) Total consumption Q is some function of S and N:

$$Q = Q(S,N) \quad . \tag{1}$$

(6) Total cost of the system C_t consists of two parts. System cost C_s is a function of S and N:

$$C_s = C_s(S,N) \quad . \tag{2}$$

Operating cost C_o, the cost of other variable inputs, is a function only of total demand for the service:

$$C_o = C_o(Q) \quad . \tag{3}$$

Total cost then is:

$$C_t = C_s + C_o \quad . \tag{4}$$

(7) The spatial pattern of facilities for a given S,N is that pattern for which Q is maximized.

(8) There is a fixed budget K for the system.

Figures 1-3 illustrate some likely properties of the expression (1)-(4). Since the service is zero priced, there is presumably some upper limit Q* to the amount that a population might be expected to consume. Holding the number of facilities constant (Figure 1), positive variation in scale, the other system variable, may be expected to produce first increasing then decreasing positive variations in demand. The curves in Figure 1 actually represent a family of sections through the surface of equation (1). They are therefore demand curves for the service, given fixed numbers of N_k of facilities at varying scales. Scale expenditures play a role of negative prices or subsidies. An exactly analogous diagram could be made for the number of facilities, holding scale constant.

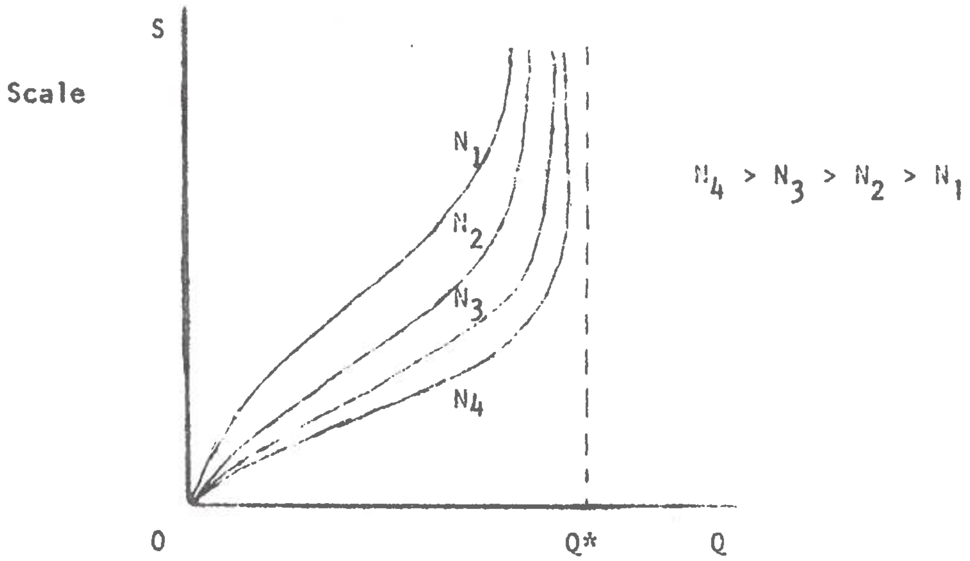

Figure 1.

The general character of Q(S,N) is thus a function monotonically increasing to some asymptote Q*. It would look like a curved surface climbing away from the origin.

Cost relationships may be handled in a similar way. Figures 2a and 2b present patterns of system cost variation for constant levels of scale and number of components, respectively. Although we assume that increases in scale eventually incur higher marginal cost, there seems to be no reason for such an increase with duplication. Rather, the reverse would seem to hold. The system cost surface, C_s, may be generated from the families of sections in Figure 2.

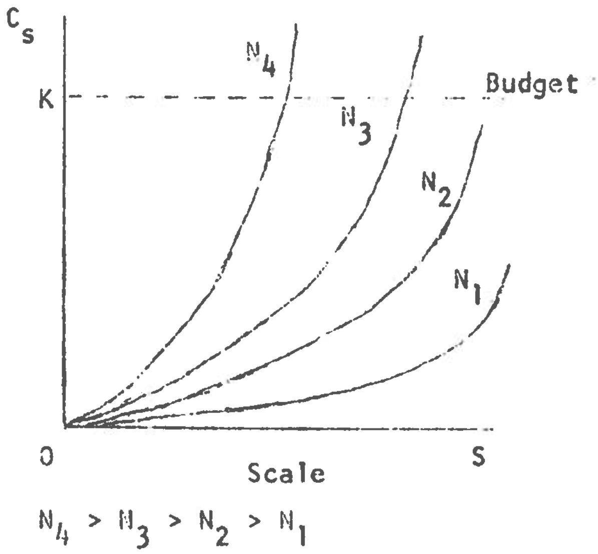

Figure 2a.

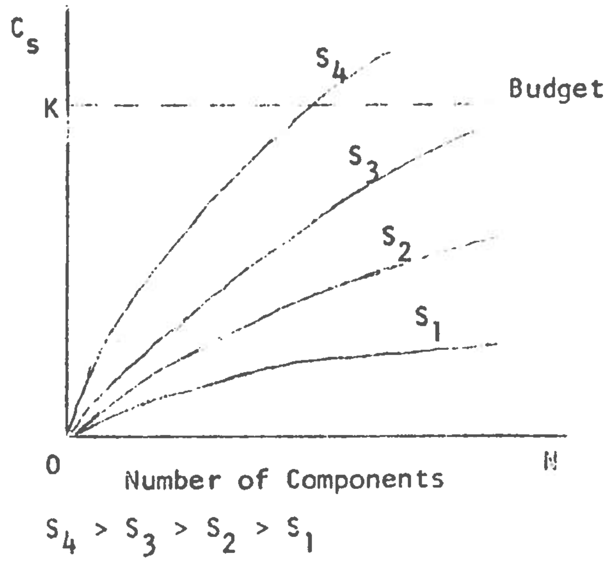

Figure 2b.

For operating cost C_o several problems arise. We have made it a function of total demand on the assumption that the marginal product of any variable input to a given system does not vary with the form of the system itself, but only with the aggregate quantity of services demanded and produced. The reason for the distinction between system and operating costs should be clear. The latter depend upon demand. They represent the variable cost of responding to demand at the level induced by the former. In part, this must be an artificial distinction. Demand for a service does respond to the level of variable inputs insofar as they determine convenience and quality of service. We will avoid this problem for the moment by assuming that variable inputs are added to maintain some constant level of quality. For simplicity this relationship is represented as generally linear in Figure 3, though it should be noted that in terms of the variables of Figures 1 and 2 it is most unlikely to be linear.

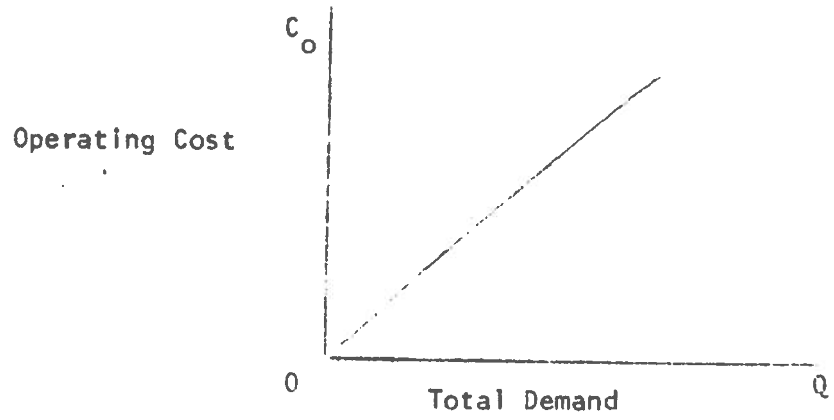

Figure 3.

With appropriate assumptions about continuity and well behaved functions, the problem may now be formulated as a constrained maximization:

$$\text{Maximize:} \qquad Q = Q(S,N) \tag{5}$$

$$\text{Subject to:} \qquad C_t = K \quad . \tag{6}$$

The Lagrangian for this problem is:

$$Z = Q(S,N) - \lambda[C_s(S,N) + C_o(Q) - K] \quad , \tag{7}$$

for which the conditions for maximization become:

$$\frac{\partial Z}{\partial S} = \frac{\partial Q}{\partial S} - \lambda\left[\frac{\partial C_s}{\partial S} + \frac{\partial C_o}{\partial Q}\,\frac{\partial Q}{\partial S}\right] = 0 \rightarrow \frac{\partial Q}{\partial S} = \frac{\lambda}{1 - \lambda\frac{\partial C_o}{\partial Q}}\,\frac{\partial C_s}{\partial S} \tag{8a}$$

$$\frac{\partial Z}{\partial N} = \frac{\partial Q}{\partial N} - \lambda\left[\frac{\partial C_s}{\partial N} + \frac{\partial C_o}{\partial Q}\,\frac{\partial Q}{\partial N}\right] = 0 \rightarrow \frac{\partial Q}{\partial N} = \frac{\lambda}{1 - \lambda\frac{\partial C_o}{\partial Q}}\,\frac{\partial C_s}{\partial N} \tag{8b}$$

$$\frac{\partial Z}{\partial \lambda} = C_s(S,N) + C_o(Q) = K \quad . \tag{8c}$$

Combining (8a) and (8b) we obtain the maximization condition:

$$\frac{\partial Q/\partial S}{\partial Q/\partial N} = \frac{\partial C_s/\partial S}{\partial C_s/\partial N} \tag{9}$$

The consequences of our assumption about variable operating cost show up immediately in (9). The equilibrium condition for demand maximization includes only system variables. If this seems peculiar, we might reflect that operating cost appears in (8c). Given our assumption that a given increase in demand generates the same operating cost no matter whether it derives from the scale or number of system components, its absence from (9) is less surprising. Whether that assumption is tenable is another matter.

More significantly, this formulation evades the problem of location via assumption (7). So far, I have been unable to incorporate the location problem into a pure analytic model. In view of the numerous heuristic approaches to this type of location problem (Cooper [4], Teitz [11]), there would seem to be advantages to structuring the total problem as a computer model. In analytical terms, this raises the problem of our assumption of continuity in the variable N. We cannot simultaneously assume it continuous for scale analysis and discrete for a locational effectiveness algorithm. Perhaps an iterative estimation process is the way around this problem, but the theoretical result is less precise. In any case, it seems probable that the location problem for public facility systems must be attacked in tandem with system structure and scale.

Several problems still remain. Introduction of variable facility scales in single system is clearly necessary. As soon as this is done, then questions of hierarchy begin to arise. We can do no more here than point out that the problem exists.

The static problem treated above is general in the sense of dealing with simultaneous location and scale of all components of a facility system. The equivalent partial problem might be formulated in several ways. If an increment to a budget for an existing system is given, then we might be interested in determining the optimal addition to the system. This does not necessarily mean that any new components are added. The entire budget increment could be spent on scale changes. If the problem is to achieve a specified incremental gain in some effectiveness measure, the same qualifications would apply. In these circumstances it is not clear how a partial form should be specified. Possibly, it should hold the present facility location structure constant and allow only scale changes and new facility locations.

4.2 Dynamic Systems

To analyze systems of facilities with static equilibrium analysis is to ignore a most important characteristic: their changes over time. Facility systems are usually built quite slowly, reacting to changes both in the size of the broader systems they serve, and in technology and social preferences. If the broader system is a growing city, then there may be conflict between static and dynamic system optima. This may be especially true if, for whatever reasons, decisions early in a system's development can effectively close off options for later forms.

A geometric illustration of a dynamic system conflict with static solutions is offered by a simple model. Consider the circular and generally symmetric city represented in Figure 4. At this particular size, for some local service the optimal number of facilities is one and it is located at the center, F_1. The city grows symmetrically both in density and at its outer margin until it reaches the size shown in Figure 5. At this new level the static equilibrium solution, taking into account a probable larger budget

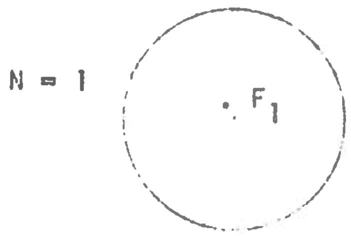

Figure 4.

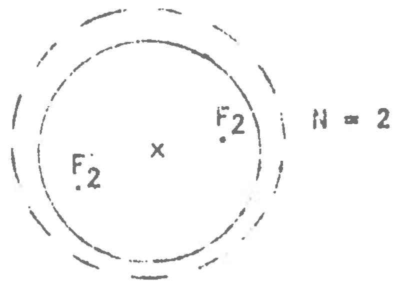

Figure 5.

for the service, calls for two identical facilities, F_2. If they are located symmetrically, there is no path of growth for this facility system from stage 1 to stage 2 that does not call for removal of F_1. Whether that is likely depends on the rate of growth and the fixed capital investment in F_1.

The example is made artificial by the assertion of identical facilities. In practice, accommodation may be partially achieved by variations in scale among facilities. For example, the equivalent problem for three components might be to approximate a symmetric uniform scale optimum (Figure 6a) by a variable scale but still symmetric three component system (Figure 6b). In the latter, the original facility is retained at a larger scale than the others. Without specifying particular forms for the relationships between spatial pattern, scale, and demand, we cannot say much more than this.

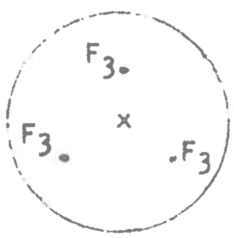

Figure 6a.

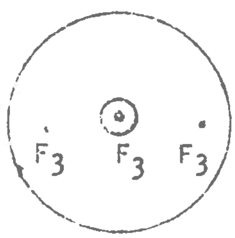

Figure 6b.

The dynamic problem suggests two lines of approach. We may look for possible system growth paths through time under varying constraints and criteria for effectiveness, and try to identify stages at which such paths coincide with static equilibrium solutions. Or we might set up static optimal solutions and try to construct minimum cost paths to connect them. Since most facility system analyses are likely to start with an existing set of components, most of which incur high relocation costs, either form could be employed. The choice is perhaps yet another version of the process-endstate conflict in planning models, in this case with both forms involving specific criteria for choice since the decisions are public.

Very little work in this direction has been done. A recent attempt by Rorick [8] to model alternative growth paths of city structure is highly suggestive but incorporates no cost constraints. Again, computer models seem to be most promising given the mathematical complexity of any reasonable looking structure for analysis.

5. Conclusion

The chief aim of this paper has been to make a case for a field of location theory specifically devoted to urban public facilities. Such an endeavor appears to be worthwhile both for its theoretic interest and its probable utility. Theoretical problems derive chiefly from the system character of the unit of location and the public decision aspect of objective functions and constraints. Practical applications are suggested by recent attempts to improve the quality of local resource allocation through such devices as PPB.

For the most part, we have catalogued problems. Several important ones remain unconsidered. Foremost among them is the question of multiple jurisdictions providing the same services within a larger area. This can occur in two ways. Multiple governments exert jurisdiction over spatially disjoint territories within most larger urban areas. The allocation of resources under such circumstances has been examined by Tiebout [12] and Williams [14]. Within any government, there is typically functional overlap between agencies or departments. This problem has been discussed chiefly in a non-theoretical way in the PPB literature. Both are likely to influence facility system location.

Finally, we should be well aware that public decisions are made in the political system. Political variables will enter any decision on a facility system, especially since public facilities are so often the visible symbols of delivery of public goods to particular groups in the city. We have not considered political variables here, being more concerned with an efficiency approach, but their power may be attested by any official who has ever tried to close an existing facility.

REFERENCES

[1] Alexander, C., 'The Pattern of Streets,' Journal of the American Institute of Planners, 32 (1966), 273-278.

[2] Berry, B.J.L., Geography of Market Centers and Retail Distribution Englewood Cliffs, N.J.: Prentice-Hall, 1967.

[3] Berry, B.J.L., and A. Pred, Central Place Studies: A Bibliography of Theory and Applications, Philadelphia: Regional Science Research Institute, 1961.

[4] Cooper, L., "Solutions of Generalized Locational Equilibrium Models,' Journal of Regional Science, 7 (1967), 1-18.

[5] Haggett, P., Locational Analysis in Human Geography, New York: St. Martin's Press, 1966.

[6] Isard, W., Location and Space Economy, Cambridge: Massachusetts Institute of Technology Press, 1956.

[7] Isard, W., et al., Methods of Regional Analysis, New York: John Wiley, 1960.

[8] Rorick, H., "Conflicts in Growth and Optimization: A Study for the Factor of Accessibility," unpublished M.Arch Thesis, University of California, Berkeley, 1967.

[9] Schaller, H., Public Expenditure Decisions in the Urban Community, Baltimore: Johns Hopkins Press, 1963.

[10] Schneider, J. B., Measuring the Locational Efficiency of the Urban Hospital, Discussion Paper Series No. 11; Philadelphia: Regional Science Research Institute, 1967.

[11] Teitz, M. B., and P. Bart, Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph, Working Paper No. 63; Berkeley, Calif.: Center for Planning and Development Research, University of California, Berkeley, 1967.

[12] Tiebout, C. M., "A Pure Theory of Local Expenditures," Journal of Political Economy, 64 (1956), 416-424.

[13] Tiebout, C. M., "Economic Theory of Fiscal Decentralization," in Public Finances: Needs, Sources and Utilization, National Bureau of Economic Research; Princeton: Princeton University Press, 1961.

[14] Williams, A., "The Optimal Provision of Public Goods in a System of Local Government," Journal of Political Economy, 74 (1966), 18-33.