Theory. Journal of Political Economy, Vol. 74, April 1966.
6. C. F. Manski. The Analysis of Qualitative Choice. Department of Economics, MIT, Cambridge, PhD dissertation, June 1973.
7. D. McFadden. Conditional Logit Analysis of Qualitative Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1974.
8. M. E. Ben-Akiva and S. R. Lerman. Disaggregate Travel and Mobility Choice Models and Measures of Accessibility. In Behavioral Travel Modelling (D. A. Hensher and P. R. Stopher, eds.), Croom-helm, London, in preparation.
9. T. J. Atherton. Approaches for Transferring Disaggregate Travel Demand Models. Department of Civil Engineering, MIT, Cambridge, MS thesis, Feb. 1975.
10. T. J. Atherton and M. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. TRB, Transportation Research Record 610, 1976, pp. 12-18.
11. W. Pecknold and J. Suhrbier. Tests of Transferability and Validation of Disaggregate Behavioral Demand Models for Evaluating the Energy Conservation Potential of Alternative Transportation Policies in Nine U.S. Cities. Cambridge Systematics and Office of Energy Conservation Policy, Federal Energy Administration, April 1977.
12. S. R. Lerman. A Disaggregate Behavioral Model of Urban Mobility Decisions. Department of Civil Engineering, MIT, Cambridge, PhD dissertation, June 1975.
13. Task Termination Report—Internal CBD Travel Demand Modeling. Barton-Aschman Associates and Cambridge Systematics, Aug. 1976.

Publication of this paper sponsored by Committee on Urban Activity Systems.

Optimizing Urban Mass Transit Systems: A General Model

Alan Black, Graduate Program in Planning, University of Texas at Austin

This paper describes a model for determining the general dimensions of an optimal mass transit system for an idealized urban area. The model is based on a circular city with a definite center and with density declining uniformly from the center in all directions according to the negative exponential function. The transit system consists of radial routes that emanate from the center and contain discrete stops. Only trips to or from the center are considered, and travel is assumed to occur only in radial and circumferential directions. The model represents total community costs of the system, defined to include travel time, operating costs, equipment, and construction. A recursive procedure was devised to find a simultaneous minimum with respect to the spacing of routes, number and spacing of stops on each route, and average headway. Numerical analyses were conducted for six hypothetical cities by using varying values for the parameters of the density function. In each case, three types of transit systems were compared: conventional bus service, buses on exclusive lanes, and rail rapid transit. The optimal system in the largest city examined was exclusive bus lanes; in the other five cases, the optimal system was conventional bus service. Other interesting relations that appeared in the results are summarized.

The United States has entered a new era of massive investment in urban mass transit, prompted by the willingness of Congress to authorize billions of dollars in federal aid for local transit systems. However, there is yet no systematic procedure for allocating these resources and determining whether a transit proposal is worthwhile. Each proposal is evaluated on an ad hoc basis, and considerable weight is given to the zeal of the proponents, political pressures, and the current availability of funds. Choice of technology has become a major issue in many areas, and the question of whether medium-sized cities should proceed with huge investments in fixed-guideway transit systems is particularly controversial.

This paper summarizes a dissertation aimed at determining the dimensions of an optimal transit system for an idealized urban area (1). The approach was to hypothesize a circular city with a definite center and with density declining uniformly from the center in all directions. The transit system consists of routes that emanate from the center and contain discrete stops. By use of integral calculus, a model was derived that represented the total community costs of building and using such a system. By use of differential calculus, a procedure was developed to optimize the principal design variables in the system: the number of radial routes, their length, and the number and spacing of stops on each route. Numerical analyses compared three common forms of conventional transit: buses on city streets, buses on exclusive lanes, and rail rapid transit.

Such an abstract model cannot be mechanically applied to the complex, irregular pattern of a real city. Abstraction is an unavoidable compromise if a model is to be made mathematically tractable. Similar approaches have been followed in many previous studies of transit optimization. A few of these will be cited here; a fuller review can be found elsewhere (1).

Most previous studies can be divided into two geometrical approaches. One of these assumes a gridiron network of transit routes laid on a homogeneous infinite city, usually with a uniform density of trip ends. What may have been the first study of this type was done by Creighton and others (2) and involved both highway and transit grids; the object was to find the optimal combination of investment in the two modes. Holroyd (3) assumed a single grid of bus routes and derived a solution for the optimal spacing of routes and frequency of service. Two dissertations, one by Mattzie at Carnegie-Mellon (4) and the other by Woodhull at Rensselaer Polytechnic (5), also dealt with grid systems of transit routes.

The second approach is to examine a single transit line. Often one terminal is assumed to be in the central business district (CBD), and only trips to or from this terminal are considered. The object is usually to find the optimal spacing of stops. Some studies have assumed uniform spacing, but the more interesting have allowed for variable spacings. An early study of variable spacing of stops was made by Schneider (6). Vuchic (7) later did a fuller analysis of the problem. Both researchers assumed a constant density of boarding passengers along the line and came to the conclusion that interstation spacings should increase as one approaches the CBD. This is opposite to conditions normally found on radial rail routes in which the spacings become smaller near the CBD.

A third geometrical approach was taken by Byrne (8), who assumed a circular city of given radius and, for the case in which population density varies only with distance from the center, derived a solution for the optimal number of radial routes. Byrne presented results for four density functions including the negative exponential. The model described here has a similar geometry and also uses the negative exponential function, but it involves simultaneously solving for the optimal number of radials and optimal length of radials as well as other variables.

DESCRIPTION OF THE MODEL

The hypothetical city is a complete circle, uniform throughout its 360 degrees, uninterrupted by barriers or irregularities, and extending to infinity. The city has a center that is taken to represent the CBD. The transit network consists of an unknown number of radial lines that emanate from the center and extend an unknown distance. Each line has discrete stops, and access is possible only at these fixed points, which must be determined. Because of the assumed symmetry of the city, the radial lines are equally spaced. Each has the same number of stops, spaced in the same way, and is of the same length.

The transit system serves a given amount and distribution of travel demand. Demand is highest at the center and declines with increasing distance from the center. Demand is assumed to be constant (i.e., land use, trip generation, and modal split are held constant). Every trip must be made, and the only alternative to using transit is walking.

To make the problem mathematically manageable, only trips to or from the CBD (assumed to be trips to or from the point at the center) were considered. In most cities, such CBD-oriented trips represent roughly half of all transit trips. This is the largest market for transit, and it dominates the design of transit networks.

It was assumed that travel can occur only in radial or circumferential directions but that otherwise it can occur anywhere on the city's surface. There are no circumferential transit routes; all circumferential travel is on foot. Each inbound traveler starts from his or her origin and walks in a circumferential arc to the nearest radial transit route. There the traveler has a choice between walking inward or outward until reaching a stop. He or she chooses the stop that minimizes the total time from origin to destination. Close to the center, some travelers find it faster to walk all the way and not use the transit service.

It was assumed that inbound and outbound trips are equal in number and form mirror images of each other. On each radial transit route, vehicles shuttle back and forth between the center and outer terminal, stopping at all designated stops.

Objective

The objective selected was to minimize total community costs, measured in dollars. Total costs were defined to include capital investment for guideway and vehicles, operating costs, and the door-to-door travel time of travelers. It would be desirable to include other types of community costs, such as externalities and intangibles, but they were omitted because of the difficulty of measuring them or converting them to dollar values.

The approach to measuring time costs was to calculate the distance traveled from door to door and to estimate average speed on the portions of the route traversed. There are several components in door-to-door travel time. One of these is the time spent in the transit vehicle, which must be divided into two parts. The first is the time that would accrue if the vehicle moved at its cruising speed from the rider's point of embarkation to the point of debarkation. The second component consists of the additional time penalties incurred when the vehicle accelerates, decelerates, and waits at stops to load and discharge other passengers.

Two other components of door-to-door travel time were included. One is walking time—the time it takes a traveler to walk from origin to boarding stop or from where he or she gets off to his or her final destination. The other is waiting time, which depends on the scheduling of vehicles. It was assumed that average waiting time is half the scheduled headway (the time between successive buses or trains).

Operating costs depend on a number of factors, but here they were based solely on vehicle kilometers. This appears to be the most significant relation and also the simplest.

There are two major types of capital costs: fixed facilities (such as roadbed, structures, and stations) and running equipment (buses and trains). In analyzing a specific proposal, detailed estimates of capital costs are based on engineering drawings. This cannot be done for a hypothetical city; hence, the cost of fixed facilities was based on kilometers of guideway and number of stations. Equipment costs were estimated on the basis of the number of buses or train cars required to serve peak-period demand plus an allowance for vehicles out of service.

The daily time and operating costs and the one-time investment costs were put on a comparable basis through the annual cost method although it is actually average weekday costs that are represented in the model. For capital costs, an expected life span and interest rate were assumed, and the equivalent annual cost was calculated. This was converted to average weekday cost by assuming a number of weekday equivalents for a year.

Decision Variables

The most important decision variables in designing such a system are

1. The number of radial routes (N),
2. The number of stops on each radial route (z),
3. The length of each radial route (x_z), and
4. The spacings between stops on each route.

The spacings between stops are implied in the set of variables x_1, x_2, x_3, ... , x_z, where x_i is the distance from the city center to the ith stop. Thus, the spacing between the second and third stops is given by $x_3 - x_2$.

Another variable—scheduling of service—is also within the control of the transit authority. There is an important relation between frequency of service and route spacing, and they should be optimized simultaneously. Frequency of service was represented by average headway over the full day (h).

Role of the Density Function

To develop the model, one must know the locations of the trip ends on the surface of the city. Each trip has one end at the center, but the other end is elsewhere. The approach was to adopt a function that relates the density of these outer trip ends to distance from the city center.

There remained the question of what function to use. A considerable body of literature, starting with the landmark article by Clark (9), suggests that the negative exponential function represents the relation between population density and distance from the city center. However, few studies have dealt with the density of trip ends. Furthermore, a review of the literature—described elsewhere (10)—revealed that there is now much debate over the proper function for population density. Researchers have used a variety of equations and obtained close fits to empirical data with many of them.

Therefore, some empirical research was conducted by using population and travel data from transportation studies of 12 metropolitan areas of the United States that range in size from New York to Syracuse. Regression analysis was used to fit data on six population and trip-end variables to four alternative equations: linear, exponential, power curve, and normal curve. The findings are described elsewhere (11).

Of particular concern was the density of CBD-oriented transit trip ends. Data for this variable were available only for six medium-sized cities. The analysis showed that the exponential function yielded the highest correlation coefficient for four of the cities. For the other two cities, the normal curve gave the highest correlation but the exponential function was almost as good. In all cases, the exponential function had a high correlation that ranged from 0.880 to 0.993.

As a consequence, the negative exponential function was selected for inclusion in the model. The specific equation is

$$D(r) = Ae^{-br} \tag{1}$$

where

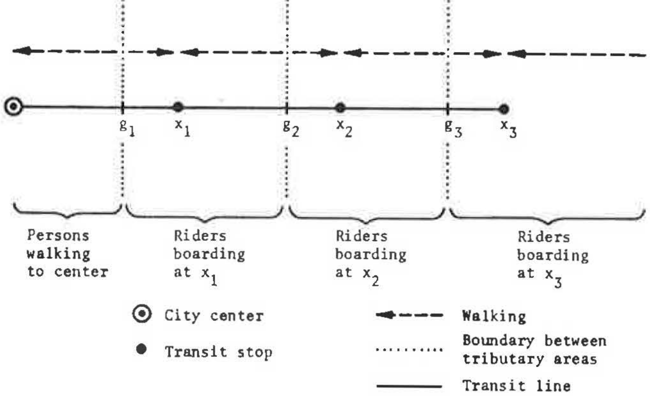
$D(r)$ = density of trip ends at distance r,
r = distance from the center,
e = base of natural logarithms, and
A and b = parameters.

DERIVATION OF THE MODEL

The model consists of a single equation, derived by integral calculus, that represents the total community costs of the transit system. Deriving the equation involves summing up the kilometers traveled, minutes spent in travel, and costs for the entire city. Because of the assumed circular nature of the city, the problem lends itself to the polar coordinate system rather than the Cartesion coordinate system. In the polar system, any point r, θ is identified by its distance from the origin r and the angle θ between a ray from the origin and a given axis.

Each stop on a route draws travelers from a tributary area that can be drawn on a map. There is also a circular area around the center from which people walk to the center and do not use transit. Therefore a route with z stops serves a sector that is divided into $z + 1$ tributary areas. This is shown in Figure 1 for a route with only three stops (plus the CBD terminal).

Figure 1. Relation between tributary areas and locations of transit stops.



The boundary between the tributary areas of adjacent stops must be determined by finding the point at which a traveler is indifferent; that is, the total travel time to the CBD is equal whichever stop is used. This calculation is incorporated into the model. The variable g_i represents the distance from the center to the boundary between those who walk in to stop $i - 1$ and those who walk out to stop i.

Describing the derivation of the entire equation would be impossible in this space. Therefore, the approach is illustrated by deriving the person kilometers traveled on transit vehicles. This is done for inbound travelers for a system with three stops, as shown in Figure 1. This requires determining the number of passengers who board at each stop and multiplying by the distance from the stop to the center. For the first stop, it is necessary to integrate the density function between the inner and outer boundaries of the tributary area (g_1 and g_2) and to multiply this by x_1. The resulting derivation is

$$\int_0^{2\pi}\int_{g_1}^{g_2} Ae^{-br}\, x_1\, r\,dr\,d\theta = (2\pi A x_1/b^2)\,[e^{-bg_1}(1 + bg_1) - e^{-bg_2}(1 + bg_2)] \tag{2}$$

Notice that this is integrated over the full circle; it represents the travel by all persons who board at the first stop on all radials. Since the city is radially symmetrical, the behavior on each radial is identical, and there is no point in calculating them separately.

For passengers who board at the second stop, the only differences are that the limits of the tributary area are g_2 and g_3 and each passenger rides x_2. Hence, the expression for person kilometers has the same form as Equation 2, or

$$\int_0^{2\pi}\int_{g_2}^{g_3} Ae^{-br}\, x_2\, r\,dr\,d\theta = (2\pi A x_2/b^2)\,[e^{-bg_2}(1 + bg_2) - e^{-bg_3}(1 + bg_3)] \tag{3}$$

For persons who board at the third stop, the expression is

$$\int_0^{2\pi}\int_{g_3}^{\infty} Ae^{-br}\, x_3\, r\,dr\,d\theta = (2\pi A x_3/b^2)\, e^{-bg_3}(1 + bg_3) \tag{4}$$

Adding these together, we get the total person kilometers traveled by transit:

$$(2\pi A/b^2)\,[e^{-bg_1}(1 + bg_1)(x_1) + e^{-bg_2}(1 + bg_2)(x_2 - x_1) + e^{-bg_3}(1 + bg_3)(x_3 - x_2)] \tag{5}$$

This is for a system with just three stops on each radial. For a larger system, there would be more terms since it is necessary to make a separate derivation for each stop. However, all the expressions are identical in mathematical form; only the designation of the variables must be altered. Thus, the equation for a three-stop system can be extrapolated to a system with more stops.

The derivation proceeded in a step-by-step (or stop-by-stop) fashion. The most difficult part was determining the kilometers walked in a radial direction. This had to be done separately for those who walk inward and those who walk outward. For a system with z stops, it was necessary to derive expressions for 2z + 1 separate areas. But again, all the expressions turned out to be identical in mathematical form.

Kilometers traveled was converted to person minutes of travel time by applying appropriate values of average speed. Waiting time and time for delays from stops were added to obtain total person minutes of travel time. This was converted to dollars by applying the assumed monetary value of travel time. Then expressions for operating, construction, and equipment costs were derived and added to get total community costs.

When the expressions for all cost components are added together, a considerable amount of cancellation and simplification is possible. The total cost equation for a system with three stops is as follows (certain portions of the calculations given in this paper were done in U.S. customary units, and in these instances no SI units are given):

$$\begin{aligned} y = \{&(4\pi A c_2 t/b^3)\,[1 + e^{-bx_1}(2 + bx_1) + e^{-bx_2}(2 + bx_2) \\ &+ e^{-bx_3}(2 + bx_3) - e^{-bg_1}(2 + bg_1) \\ &- e^{-bg_2}(2 + bg_2) - e^{-bg_3}(2 + bg_3)]\} \\ &+ (2\pi^2 A c_2 t/N b^3) + (2NKqx_3/h) \\ &+ [(2mNV/ph)(c_1 x_3 + 3d + L)] + (INx_3 + 3JN) \end{aligned} \qquad (6)$$

where

- y = total costs,
- x_i = distance from center to ith stop (km),
- g_i = distance from center to boundary between tributary areas for stop $i - 1$ and stop i (km),
- c_2 = walking speed (min/km),
- t = value of travel time (dollars/min),
- N = number of radial routes,
- K = length of daily transit service period (min),
- q = operating cost per vehicle mile (dollars),
- h = headway between buses or trains (min),
- m = spare vehicle factor,
- V = equivalent daily cost of a vehicle (dollars),
- p = ratio of peak-period headway to average all-day headway,
- c_1 = cruising speed of transit vehicle (min/km),
- d = delay for a stop (min),
- L = layover time (min),
- I = equivalent daily cost of a mile of guideway (dollars), and
- J = equivalent daily cost of a station (dollars).

The right-hand side of Equation 6 is divided into five parts, each of which has a recognizable significance, so that the equation can be rewritten in verbal form as follows: Total costs = radial travel, delay, and waiting time + circumferential travel time + operating cost + equipment cost + construction cost.

FINDING THE OPTIMAL SOLUTION

The total cost equation for a system with z stops on each radial has z + 2 decision variables. These are N, h, and the set of z variables that represent the distance from the center to each of the stops (x_1, x_2, x_3, . . . , x_z). x_z is also the length of the route.

To find a global minimum for all the variables, one starts by taking the partial derivative of the equation with respect to each variable and setting each result equal to zero. This yields a set of z + 2 equations that contain z + 2 unknowns. The next step is to find a simultaneous solution for the z + 2 equations that will specify the optimal solution. But all of the equations are nonlinear, and no general analytical method exists for finding the simultaneous solution to a set of nonlinear equations. Consequently, an approximating procedure was developed to find the optimal solution.

The set of equations does have a special structure that can be exploited to develop a recursive procedure. Most of the equations contain three unknowns, but there is one that contains only two. The equations can be sequenced so that, by assuming a value for one variable, values can be calculated for all the other variables. When this has been done, one equation remains. Inserting the previously calculated values in this equation provides a check on whether a simultaneous solution has been obtained.

Usually a simultaneous solution will not occur at first because the process began by guessing the value for one variable. Therefore, the recursive procedure must be embedded in a search procedure to find the right starting value. A classical one-dimensional search technique known as the regula falsi method was adopted for this purpose. This involves making two trials with arbitrarily selected values, comparing the results, and calculating a "best guess" for a third trial. The method continues with successive trials, always comparing the results of the two previous trials, until convergence is reached. The recursive procedure and the regula falsi method were implemented in a computer program that proved efficient in approximating a simultaneous solution for the set of equations.

There is a further, unusual dimension to the problem. To carry out the procedure just described, one must specify the number of stops because this determines the number of unknowns and equations. However, the number of stops is itself an unknown of some importance. Thus, at the start, the number of equations to be solved is unknown! This problem was handled by embedding the above procedure in an overall search for the optimal number of stops. This was done by using the Fibonacci search method, which successively eliminates groups of integers until it locates the integer that gives the optimum.

NUMERICAL ANALYSIS

The computer program was used to calculate optimal transit systems for a number of hypothetical cases. Two interests were of primary importance in the selection of the tests:

1. Choice of technology—What is the best transit mode for a particular city? Tests were run for the three best known types of conventional transit service: (a) ordinary bus service in which buses run in mixed traffic on surface streets, (b) exclusive bus lanes, and (c) rail rapid transit. These are referred to as local bus, busway, and rail.
2. Impact of the density profile—What effect does the density of trip ends have on the optimal transit system? Tests were made for six hypothetical cities with different values for parameters of the density function. The three transit modes were compared for each city so that optimal transit systems were calculated for 18 cases.

The table below gives the regression and correlation coefficients obtained for the exponential function for six metropolitan areas:

City	A	-b	-r
Detroit	3427	0.286	0.966
Cleveland	4043	0.275	0.992
Pittsburgh	2345	0.355	0.880
Buffalo	2145	0.384	0.974
Rochester	2705	0.724	0.993
Syracuse	1285	0.632	0.935

The value of A is the density of trip ends per square mile at the city center (these numbers represent only the outer ends of trips, which is why they seem low). The value of b indicates the rate of decline in the density of trip ends per square mile with each mile of increased distance from the city center. The higher the absolute value of b, the more compact the city.

It was decided to make tests with combinations of two A values (2000 and 4000) and three b values (0.25, 0.50, and 0.75). Each city is given a "name" that signifies the two parameter values: for example, city 4/25 is the case where A = 4000 and b = 0.25. The total number of person trips made in each city can be determined from the following result:

$$\int_0^{2\pi}\int_0^{\infty} Ae^{-br}\, r\, dr\, d\theta = (2\pi A/b^2) \tag{7}$$

Table 1. Dimensions of optimal local bus systems.

City	Number of Stops	Length of Each Route (km)	Average Spacing (km)	Number of Radials	Total Route (km)	Headway (min)
4/25	19	25.42	1.34	60	1526.2	8.60
2/25	19	25.43	1.34	43	1082.8	12.20
4/50	13	12.54	0.96	22	273.6	12.50
2/50	13	12.55	0.97	16	195.0	17.80
4/75	10	8.27	0.83	12	101.3	15.80
2/75	9	7.72	0.86	10	75.6	20.60

Note: 1 km = 0.62 mile.

The person-trip totals for the six hypothetical cities are given below:

City	Person Trips
4/25	402 124
2/25	201 062
4/50	100 531
2/50	50 265
4/75	44 680
2/75	22 340

Numerical analyses also require specifying values for a large number of parameters that mostly represent cost and performance characteristics of the transit system. The literature was surveyed to ascertain reasonable values for all parameters. Many of these depend on the transit mode being analyzed, but some (such as the interest rate—assumed to be 10 percent—and the value of travel time—set at \$2.40/h) are common to all modes.

Local Bus System

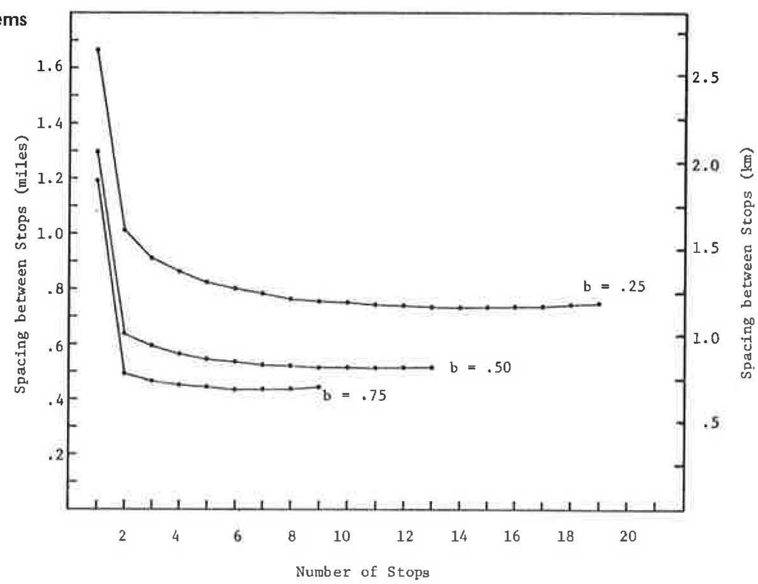
Dimensions of the optimal systems for the six cities are given in Table 1. As one would expect, the larger the city is, the larger is the optimal transit system. The optimal number of stops ranges from 9 to 19; the optimal length of each route from 7.72 to 25.43 km (4.8 to 15.8 miles); and the optimal number of radials from 10 to 60. Thus, there is considerable variation in the optimal values, which clearly depend on the density profile.

The value of A has little effect on the optimal number and spacing of stops or the optimal route length. The value of b has much more influence: The more compact the city is, the shorter are the routes, the fewer are the stops, and the closer is the spacing between stops.

Where A does have an impact is on the number of radials (although the b parameter remains dominant). City 4/25 has twice as many trips as city 2/25, and the optimal number of radials increases from 43 to 60. This suggests that the response to a uniformly distributed increase in demand should be to increase the number of routes rather than the frequency of stops on existing routes.

The optimal value for average headway ranges from 8.6 to 20.6 min; it is influenced by both the A and b values. When the value of A is doubled, optimal headway

Figure 2. Spacing between stops for optimal local bus systems for cities with A = 2000.



is reduced by less than half. This suggests that, when bus trips are added to handle increased demand, they should be distributed between existing routes and newly created routes (which will reduce both walking and waiting).

The pattern of spacing between stops is also of interest. The same pattern was found in all cases and persisted in the busway and rail alternatives. To illustrate this pattern, the interstop spacings for the three cities with Λ = 2000 are shown graphically in Figure 2. The pattern has the following features:

1. Starting from the center, the spacing decreases outward to a point about four-fifths the length of the route.
2. From this point to the outer terminal, the spacing gradually increases.
3. The first stop has a much larger spacing than any other, but variation among the others is very slight. One could generalize by saying that, except for the first stop, the spacing should be approximately uniform. This seems more realistic than the optimal spacing pattern derived by Schneider (6) and Vuchic (7).

Busway System

The principal difference in the busway system is that it involves construction costs for guideway and stations in return for which the buses achieve higher speeds and lower operating costs. Table 2 gives the dimensions of the optimal busway systems. These are substantially smaller than the local bus systems for the corresponding cities in terms of number of stops, number of radials, and total kilometers of route.

The average spacing between stops is greater in all cases than for local bus. This is a concomitant of faster bus speeds: the faster the speed of transit service is, the farther people will walk to use it. The delay for a stop also increases, which further increases the spacing.

The value of A has more impact on the number of stops and on route length than it does in the case of local bus. This is undoubtedly because each station and kilometer of route entails a construction cost. Doubling the number of trips distributes this capital cost more and justifies a higher level of investment.

Optimal headways are much lower for busway than for local bus. This is apparently because the busway involves high costs for stations and route kilometers (versus zero cost for local bus) but the operating cost per vehicle kilometer is much lower. The outcome is a smaller route structure with better service.

Table 2. Dimensions of optimal busway systems.

City	Number of Stops	Length of Each Route (km)	Average Spacing (km)	Number of Radials	Total Route (km)	Headway (min)
4/25	13	24.48	1.88	16	401.6	2.35
2/25	11	22.64	2.06	12	272.7	3.45
4/50	7	10.19	1.46	9	87.1	4.90
2/50	5	8.75	1.75	6	56.5	7.40
4/75	4	5.64	1.41	6	33.3	7.60
2/75	3	4.37	1.46	5	21.0	8.50

Note: 1 km = 0.62 mile.

Rail System

The principal difference between the busway and rail alternatives is that rail is more capital intensive. For rail, the costs of building a station and a route kilometer were assumed to be twice as great. The cost of a railcar was assumed to be more than four times greater than the cost of a bus.

Table 3 shows the dimensions of the optimal rail systems. Each is smaller than the optimal busway system for the same city. The number of stops and the length and the number of radials are all reduced. The slightly greater average spacing between stations results from an assumed higher cruising speed and larger delay for a stop with the rail alternative.

The rail model was slightly different from the one used for the bus alternatives in that it included a feature that also optimized the average length of trains. The results are given in the last column of Table 3, which shows that, in cities with greater total demand, trains should be longer. This in itself is not surprising, but note that optimal headway is not greatly reduced in the larger cities. This suggests that greater demand should be handled by running longer, rather than more frequent, trains.

Comparison of the Three Modes

The busway had the least total cost for the largest city examined—city 4/25. The local bus system had optimal cost for the other five cities. The rail system had the highest cost in all cases.

The computer program calculates many other characteristics of the optimal systems. It is of interest that in two cases the busway system had the highest average travel speed from origin to destination, whereas the local bus system was highest in the other four. The rail alternative turned out poorly in this regard because walking was assumed to be the only mode of access, and walking distances were quite high. This suggests the importance of supplementing rail lines with feeder bus service.

The optimal headways are of interest because increasing the frequency of service is a common policy objective. The busway system was best in all cities, and rail was uniformly second. This again demonstrates that when there is a construction cost the optimum produces a small route structure with frequent service. When there is no construction cost, the route structure is much larger, and service on each route is less frequent.

Table 3. Dimensions of optimal rail systems.

City	Number of Stops	Length of Each Route (km)	Average Spacing (km)	Number of Radials	Total Route (km)	Headway (min)	Average Number of Cars per Train
4/25	11	23.85	2.17	11	265.0	8.55	3.58
2/25	9	21.17	2.35	8	179.3	10.00	2.75
4/50	5	8.64	1.73	6	55.7	8.00	1.45
2/50	4	6.97	1.74	5	35.8	9.05	1.03
4/75	3	4.40	1.47	5	20.9	9.20	1.00
2/75	1	1.93	1.93	5	9.7	9.70	1.00

Note: 1 km = 0.62 mile.

CONCLUSIONS

The model appears to give a reasonable representation of total costs for different types of transit systems in cities with different density profiles. The results indicate considerable sensitivity to the form of transit service and the parameters of the density function. There are weaknesses in the current formulation; improvements and extensions are certainly possible. It would be desirable to make transit demand (the number of trips) and land-use configuration (implied by the density profile) sensitive to the provision of transit service, to include trips not going to or from the CBD, and to add some type of feeder routes to the radial lines.

The study indicated that an areawide rail transit system, without supplementary conventional bus service, is less economical than an areawide busway system with the same limitation within the range of density parameters examined (roughly those of medium-sized American cities). This finding depends on the values assumed for the cost and performance parameters, especially construction cost. Some medium-sized cities may contain sectors that have atypically high densities that would justify a rail line. There also may be situations in which alignments can be obtained at unusually low cost—perhaps underused railroad rights-of-way or the median strips of freeways. Any situation that involves atypically low costs for land acquisition and construction is more likely to warrant a rail line. This also applies to the busway system, which proved more expensive than conventional bus service in five of the six hypothetical cities.

It is surprising that ordinary bus service did so well in the comparison. This was largely because a dense network and close spacing of stops produced substantially shorter walking distances than did the alternatives. This underlines the importance of complementing high-speed main-line facilities with a pervasive feeder system or parking facilities at stations or both.

Both the A and b parameters of the density function affect the optimal transit system, but the b parameter has much more influence. Its major impact is on the length of radial routes and the number of stops. The average spacing between stops did not vary much from city to city.

Factors that vary in response to the A parameter can be interpreted as sensitive to scale. The results indicated that some economies of scale exist, but they are not overly significant. They seem to be largest when construction costs are involved.

Historical evidence is conclusive that the values of A and b have been declining in cities all over the world, which means that cities are becoming more dispersed and less centralized. It is noteworthy that the only city in which the busway alternative was optimal had the lowest value of b (city 4/25). Therefore, the historical decline of b values—although certainly related to increasing use of the automobile—does not necessarily spell doom for fixed-guideway transit systems. What happens is that trip lengths become longer, which makes it more worthwhile to introduce high-speed capital facilities.

This view conforms with the understanding of transportation planners. In most cities, the total number of trips to and from the CBD has remained fairly constant for years. However, homes are moving outward, and people are coming to the CBD from farther and farther away. This means that some radial transit improvements that could not be justified in the past may be warranted in the future.

REFERENCES

1. A. Black. Optimizing the Design of Urban Mass Transit Systems: A General Model. Cornell Univ., PhD dissertation, 1975.
2. R. L. Creighton and others. Optimum Investment in Two-Mode Transportation Systems. HRB, Highway Research Record 47, 1964, pp. 23-45.
3. E. M. Holroyd. The Optimum Bus Service: A Theoretical Model for a Large Uniform Urban Area. In Vehicular Traffic Science (L. C. Edie, R. Herman, and R. Rothery, eds.), American Elsevier, New York, 1967, pp. 308-328.
4. D. E. Mattzie. Ubiquitous Rapid Transit Service in an Urban Area. Carnegie-Mellon Univ., PhD dissertation, 1968.
5. J. Woodhull. Urban Public Transportation: Specification of Service Parameters for a Two-Mode Cooperative Hierarchy. Rensselaer Polytechnic Institute, PhD dissertation, 1970.
6. M. Schneider. Travel Time Minimization– Transit. Chicago Area Transportation Study Research News, Vol. 4, June 2, 1961, pp. 10-16.
7. V. R. Vuchic. Interstation Spacings for Line-Haul Passenger Transportation. Univ. of California, Berkeley, PhD dissertation, 1966.
8. B. F. Byrne. Public Transportation Line Position and Headways for Minimum User and System Cost in a Radial Case. Transportation Research, Vol. 9, 1975, pp. 97-102.
9. C. Clark. Urban Population Densities. Journal of Royal Statistical Society, Vol. 114, 1951, pp. 490-496.
10. A. Black. Urban Density Profiles: A Guide to the Literature. Council of Planning Librarians, Monticello, IL, Exchange Bibliography 1160, 1976.
11. A. Black. Urban Density Profiles in Mass Transit Planning. Paper presented at 22nd North American Meeting of Regional Science Association, Nov. 1975.

Publication of this paper sponsored by Committee on Urban Activity Systems.