

OPTIMIZATION AND SCALE ECONOMIES
IN URBAN BUS TRANSPORTATION

by

Herbert Mohring

Discussion Paper No. 5, June 1971

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

Optimization and Scale Economies in Urban Bus Transportation

By Herbert Mohring*

A cumulative deterioration of urban mass transportation service -- fewer riders lead to less frequent service leads to fewer riders lead to ... -- is frequently noted and deplored. A variety of rather exotic labels has been attached to this phenomenon. William Baumol [p. 425], for example, has referred to it as an example of "dynamic externalities." The appropriate social response to declining mass transit quality can, I think, more easily be seen by recognizing this phenomenon to be an example of what happens when demand declines for a commodity the production of which involves increasing returns to scale. The purposes of this note are to justify this assertion and to suggest the magnitude of mass transit scale economies and hence the lower bound¹ for an optimal transit subsidy policy.

Transportation differs from the typical commodity of price theory texts in that travelers and shippers play a producing, not just a consuming role. In using common carrier services, they must supply scarce inputs, their own time or that of the goods they ship, that are essential to the production process. In dealing with many transportation problems, it is useful to separate these two roles. It is useful, that is to say, to analyze transport costs as if user inputs are purchased in factor markets rather than supplied in kind and to treat transport demand as if the price of a trip equals whatever fare is charged plus the money value the traveler attaches to the time his trip requires.

Partially to justify these assertions, Section I develops simple optimization models for a welfare maximizing public authority which provides bus service between Here and There as well as for a monopoly bus company engaged in the same activity. Using this theoretical development as a base, Section II develops cost models for "steady state" and "feeder" bus routes. Also in this section, cost and related data approximating those which prevail in the Twin Cities metropolitan area are utilized to infer long run average and marginal costs as well as marginal cost-based user fares. An appendix goes into greater detail on the cost data employed.

I. Bus Route Optimization³

Suppose n consumers utilize the services of a bus line which provides trips between Here and There. Consumer i ($i=1, \dots, n$) derives utility from consuming a^i units per week of a general purpose commodity, dollars. He also derives utility from what happens There during each of the b^i trips per week he takes from Here to There and back. However, he incurs dis-utility from the time, $\tau^i = b^i t$, he spends traveling where t is travel time per trip -- a function, $t(B, X)$, of the total number of trips taken, $B = \sum b^i$, and of X , the total number of bus hours of service provided on the route each week.³

Consumer i 's problem, then, is to maximize his utility, $u^i(a^i, b^i, \tau^i)$ subject to his budget constraint, $I = a^i + F b^i$ where F is the fare per bus round trip. Setting up the Lagrangian expression

$$z^i = u^i(a^i, b^i, \tau^i) + \lambda^i(I^i - a^i - F b^i)$$

and differentiating with respect to $\underline{a^i}$ and $\underline{b^i}$ yields:

$$(1a) \quad z^i_{a^i} = u^i_{a^i} - \lambda^i = 0$$

$$(1b) \quad z^i_{b^i} = u^i_{b^i} + u^i_{\tau^i} t + u^i_{\tau^i} b t_B - \lambda^i F = 0$$

as first order conditions for utility maximization where subscripts refer to partial derivatives. It seems reasonable to suppose that consumer $\underline{i}$ ignores the effect his trips have on his own travel time. If so, $\underline{u^i_{\tau^i} b t_B}$ can be ignored. Dividing equation (1b) by equation (1a) would then yield:

$$(2) \quad u^i_{b^i}/u^i_{a^i} + t u^i_{\tau^i}/u^i_{a^i} = F$$

In equation (2), $u^i_{\tau^i}/u^i_{a^i}$, the ratio of the marginal disutility of travel time to the marginal utility of dollars, has the dimension dollars per hour. It therefore seems reasonable to substitute for this ratio $-\underline{v^i}$, the money cost consumer $\underline{i}$ attaches to his travel time. Doing so changes equation (2) to:

$$(3) \quad u^i_{b^i}/u^i_{a^i} = u^i_{b^i}/\lambda^i = F + \underline{v^i} t$$

This relationship says that consumer $\underline{i}$ will equate the ratio of the marginal utility of bus trips to that of dollars with the fare plus the time cost of a trip.

Suppose a Lange-Lerner bus authority wishes to maximize a welfare function, $W(u^1, \dots, u^n)$, subject to the constraint $\underline{R} = \underline{A} + \underline{C}X$ where $\underline{R}$ is the weekly flow of services available from the stock of resources at society's disposal, $\underline{A}$ is weekly consumption of dollars, $\underline{\Sigma a^i}$, and $\underline{C}$ is the resource service cost of providing a bus hour's services. Setting

up the Lagrangian expression:

$$(4) \quad Z = W(u^1, \dots, u^n) + \eta(R - A - CX)$$

and differentiating with respect to $\underline{a}^i$ and $\underline{b}^i$ yields:

$$(5) \quad W_i u^i_a - \eta = W_i \lambda^i - \eta = 0$$

$$(6) \quad W_i (u^i_b + u^i_\tau t) + \sum W_j u^j_\tau b^j t_b = 0$$

as first order conditions where $\underline{W}_i$ is $\frac{\partial W}{\partial u^i}$, the marginal welfare weight attached to individual i . The second equality in (5) follows from equation (1a). By appropriately substituting equations (2), (3), and (5), equation (6) can be shown to reduce to:

$$(7) \quad \eta(F - Vt_b) = 0$$

where $\underline{V}$ is the weighted (by number of trips taken) average value of travel time, $\frac{\sum b^i v^i}{\sum b^i}$. The Lagrangian multiplier, $\underline{\eta}$, can be interpreted as the welfare gain resulting from a one unit increase in available resource services. It is presumably positive. Hence, equation (7) can be interpreted as saying that, if welfare is to be maximized, the fare per trip must equal the difference between the marginal and the average time costs of a trip. That is, the fare must equal the additional time costs resulting from an additional trip less those time costs incurred by the trip taker himself.

Differentiating equation (4) with respect to $\underline{X}$, the number of bus hours of service provided, and making substitutions similar to those which led from equation (6) to equation (7) yields:

$$(8) \quad -\eta(VBt_x + C) = 0$$

This is precisely the same expression as would result from selecting that

value of $\underline{X}$ required to minimize $\underline{VBt(B, X) + CX}$, the total time and dollar costs of $\underline{B}$ trips.

It is of interest to compare the results summarized by equations (7) - (8) with those that would eventuate if bus service between Here and There was provided by a profit maximizing monopolist. He would presumably be interested in selecting that fare and level of service which would maximize:

$$(9) \quad \pi = FB - CX$$

where, from the standpoint of the monopolist, $\underline{B}$ can be written as

$\underline{B} = \sum b^i (F + v^i t)$. Differentiating (9) with respect to $\underline{F}$ yields:

$$(10) \quad \partial\pi / \partial F = B + F \partial B / \partial F$$

where:

$$(11) \quad \partial B / \partial F = \sum b^{i'} [1 + v^i t_{\theta} \cdot \partial B / \partial F]$$

Rearranging (11) to obtain an explicit relationship for $\partial B / \partial F$, substituting this relationship into equation (10), and again rearranging terms yields an expression for the profit maximizing fare which can be written:

$$(12) \quad F = B V^* t_{\theta} - B / B'$$

where $\underline{B'} = \sum b^{i'}$ and where $\underline{V^* = \sum v^i b^{i'} / \sum b^{i'}}$, the weighted (by the slopes of bus trip demand schedules) average value of travel time.

Rearrangements of equation (12) yields:

$$(13) \quad [(F + V^* t) - (V^* t + B V^* t_{\theta})] / (F + V^* t) = -B / [B'(F + V^* t)]$$

This is an expression quite similar to that which characterizes a profit maximizing monopolist: Lerner's index of monopoly power, $(P - MC)/P$, equals the reciprocal of the elasticity of demand for the product.

Indeed, if $\underline{V}^*$ happens to equal the trip weighted average value of travel time, $\underline{V}$, equation (13) would be identical to the profit maximizing monopoly condition. Equality of $\underline{V}^*$ and $\underline{V}$ could hold by accident. It would hold if each traveler places the same value on his travel time or if each consumer's trip demand schedule takes the form:

$$\ln b^i = c^i + k P^i$$

where $\underline{P}^i$ equals $\underline{F} + \underline{v}^i t$ and $\underline{k}$ is a factor of proportionality between $\underline{b}^{i'}$ and $\underline{b}^i$ which is the same for each consumer.

Differentiating equation (9) with respect to $\underline{X}$ and proceeding in a fashion similar to that involved in developing equation (12) yields, for the profit maximizing level of $\underline{X}$, that which satisfies:

$$(14) \quad F \partial B / \partial X = - B V^* t_x = C$$

This relationship differs from the corresponding condition for the welfare maximizing bus authority, equation (8), only in the presence of $\underline{V}^*$ rather than $\underline{V}$. Thus, holding trip consumption, $\underline{B}$, fixed, the Lange-Lerner authority would provide higher quality service (i. e., more bus hours per trip) than would the monopolist if and only if $\underline{V}^*$ is less than $\underline{V}$. There are no readily apparent a priori reasons for supposing $\underline{b}^{i'} / \underline{b}^i$ to be greater (or less) for passengers with high values of travel time than for those with low values. Hence, there are no reasons to suppose that the authority's service would necessarily be superior.

II. Scale Economies and Optimal Tolls and Subsidies in Urban Bus Transportation

Before dealing with the scale economy and related aspects of bus company operations, it seems desirable briefly to discuss the interrelationships among short and long run cost schedules and the nature of the subsidy required if short run marginal cost pricing is to be practiced by an increasing returns activity.

Suppose that producing widgets requires two inputs, labor and capital, and that the long run marginal and average cost schedules associated with this process are as drawn in Figure 1. Suppose also that the widget demand schedule (not drawn to avoid clutter) intersects the long run marginal cost schedule at $\underline{C}$. To minimize the cost of producing $\underline{OQ}$ widgets per week would require employing that amount of capital which would generate the short run marginal and average variable cost schedules shown passing through points $\underline{C}$ and $\underline{B}$ respectively. Setting price equal to long run (equals short run) marginal cost would then yield revenues of $\underline{OFCQ}$ per week. These revenues would suffice to cover the total costs of labor inputs, $\underline{OACQ}$ ($= \underline{OEBQ}$). In addition, they would yield quasi-rents of $\underline{ACF}$ ($= \underline{EBCF}$) on capital inputs. These quasi-rents would fall short of the weekly costs of capital inputs. Specifically, since this cost is $\underline{EBDG}$ ($= \underline{ACH}$), a subsidy of $\underline{FCDG}$ ($= \underline{FCH}$) would be required to cover total costs.

It would make no difference, of course, whether the widget manufacturer used his subsidy check to pay interest on his bonded indebtedness or to cover part of his wage bill. Still, for the purposes at hand, it

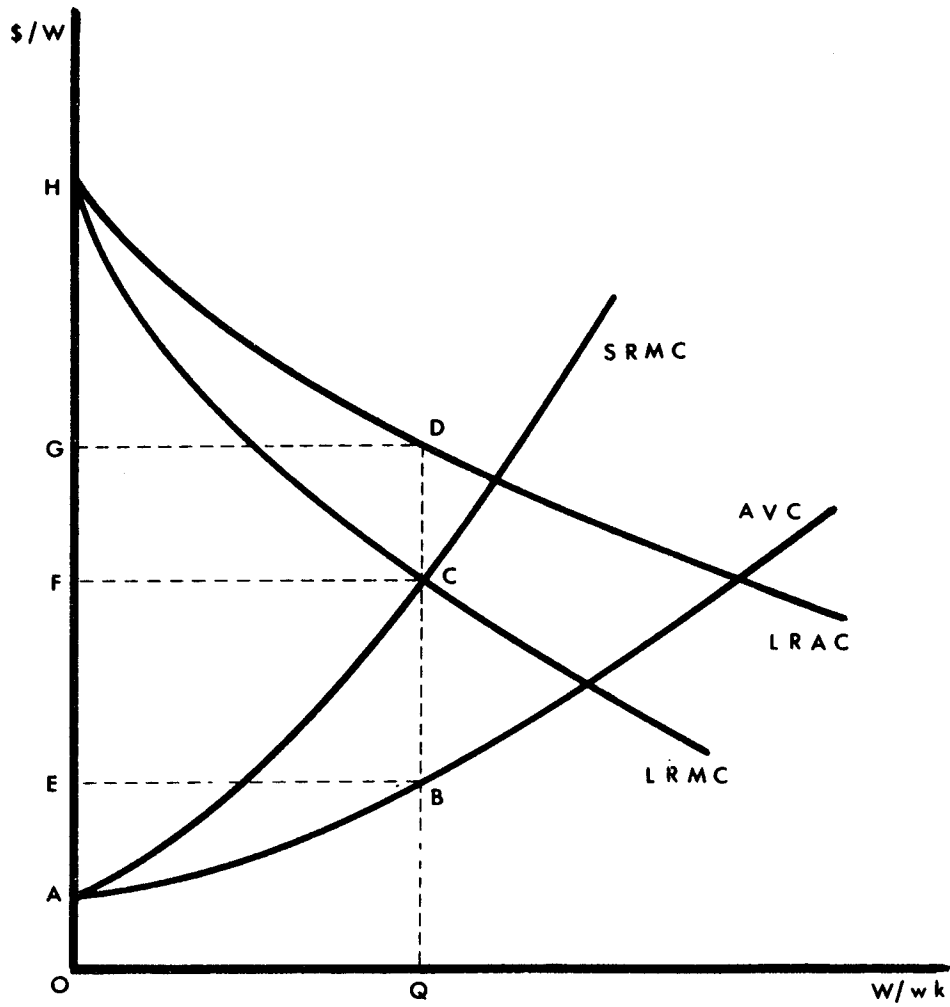


Figure 1

is important to recognize that the required subsidy equals precisely the amount by which weekly capital costs exceed the quasi-rents on capital equipment which minimum cost production and marginal cost pricing would generate.

As Section I suggests, Figure 1 can be used without alteration to describe bus company operations. The only difference in interpretation is that OEBO equals the weekly wage bill in the widget case and the weekly value of customer supplied travel time inputs in the bus case. In both cases, EBCF is a quasi-rent on capital equipment, bus or machine services as the case may be. In both cases, the required subsidy, FCDG, is the amount by which rents fall short of capital equipment costs. To the degree that scale economies exist in bus operations, then all of the subsidy required to make marginal cost pricing viable should be paid to the bus operator. Regardless of the increase in patronage that would result, it would be inefficient to employ subsidies to reduce tolls below BC, the level necessary to establish marginal cost prices for trips.

In suggesting the nature of the scale economies which arise in bus operations, it is useful to distinguish between two types of customer supplied time inputs. These are time in transit and waiting time -- time the user spends walking from his origin to a bus stop, waiting there for a bus to come, perhaps waiting at a transfer point for a second bus to arrive, walking from a bus stop to his final destination, and possibly waiting at that destination.⁴

If schedules are not published, or if published, not adhered to; if arrivals of buses at major transfer points are not synchronized; and if all trips require specific arrival times and no attempts are made to have buses reach major destinations just before common business opening hours, the average mass transit user could expect a wait of one-half the headway (i. e., the time interval) between successive buses at origin, destination and transfer points. That is, under these extreme conditions, an expected wait of $1 + a/2$ times the average bus headway would result if a percent of all riders make transfers. Most transit companies do attempt to reduce waiting time in all of the ways suggested. Unfortunately, no evidence is available on the success with which these efforts have met. A wait of half the average headway between buses (a commonly used number) will be used in all the arithmetic calculations that follow.

Suppose, for sake of illustration, that a bus company initially provides service every 20 minutes on a given route. Suppose, in addition, that the demand for service suddenly doubles and that the bus company responds by doubling the number of buses serving the route. Total bus company outlays per passenger would then remain unchanged as would the amount of time a representative passenger spends aboard a bus.⁵ However, bus headways and hence waiting time per passenger would both be cut in half if the average wait for service is, in fact, proportional to the headway between buses. Thus, the aggregate amount of time passengers spend waiting for buses to come would be the same after as before the increase

in demand for service. More generally, if bus service is provided at a rate proportional to that at which passengers travel, and if the average wait for service is proportional to the headway between buses, then total waiting time is independent of the number of passengers carried.

To put it differently, under the assumed bus line operating conditions, the gap between the average and marginal costs of a trip equals the value the average bus passenger places on the time he waits for service. Analysis of Twin City Lines schedules at a sample of points drawn a few years ago from Minneapolis and its immediately adjacent suburbs revealed distributions of elapsed times between buses with means of 9.3 and 15.6 minutes respectively during the peak and off peak demand periods.⁶ Suppose, to pick a number out of the hat, that waiting time is valued at \$1 per hour. If the average wait is half the bus headway, the gap between average and marginal costs would then equal about 8 and 15 cents respectively during peak and off peak periods. During the average 1962 weekday, the mornings and afternoon peak periods (6-9 a.m. and 3-6 p.m.) accounted for 62 percent of total Twin City Lines patronage. Hence a mean gap between average and marginal costs of roughly $0.62 \times 8\text{¢} + 0.38 \times 15\text{¢} = 10.7\text{¢}$ is suggested. The bus company's current basic adult fare is 30¢. A subsidy per passenger of 10.7 cents would therefore provide a substantial proportionate increase in the bus company's total revenues and hence its capacity to provide improved service.

Actually, assuming service frequencies to be proportional to bus line patronage understates attainable mass transit scale economies. To minimize total costs would require responding to an increase in the demand for service with a less than proportionate increase in service frequency. Both the nature of optimal service characteristics and the magnitude of bus route scale economies can be inferred from analysis of some simple bus line cost models. During most of the day on real world bus routes, about the same number of people travel in each direction along a route. During morning peaks, however, substantially more people travel toward than away from the central business district. The reverse is true of afternoon peaks. The numerical calculations described below were carried out for routes with both balanced and unbalanced flows. However, to make an already very complicated notation as simple as possible, balanced flows are assumed in the descriptions which follow of the "steady state" and "feeder" routes. Given this assumption, if service on one side of a street is optimized, optimization of service on the other side automatically follows.

Consider first a segment of a steady state route: Along each mile of the route,

$\underline{B}$ people per hour board and $\underline{B}$ exit from buses. Their origins and destinations are uniformly distributed along the route.

$\underline{M}$ is the length of each person's trip. Hence, at any point along the route segment, (approximately) $\underline{MB}/\chi$ travelers are aboard each bus where

$\underline{x}$ (to be optimized) is the number of buses that traverse the route segment each hour.

$\underline{C}$ dollars per hour is the cost of providing the services of a bus.

$\underline{Y}$ (also to be optimized) is the number of uniformly spaced bus stops per mile.

$\underline{y}$ is the speed at which travelers walk to and from bus stops.

$\underline{\beta}$ times the headway between buses is the average length of a passenger's wait for service once he reaches a stop.

$\underline{V}$ dollars is the value each passenger places on an hour spent aboard a bus while

$\underline{\alpha V}$ is the value he attaches to time spent walking to and from bus stops and waiting for buses to come. Empirical work noted in the appendix suggests $\underline{\alpha}$ to be substantially greater than 1.

$\underline{S}$ miles per hour is the overall average speed of a bus while

$\underline{S^*}$ is the speed at which it travels when not engaged in stopping and starting maneuvers.

$\underline{\epsilon}$ hours are required to board or unload a passenger once the bus has stopped and its doors have been opened.

$\underline{\delta}$ hours are added to the time required for a bus to traverse the route segment by each stopping and starting maneuver.

The total hourly costs of providing service to an $\underline{M}$ mile segment of this route can be broken into four components: bus company operating costs, and the costs to passengers of walking time, waiting time,

and time in transit. It takes M/S hours for the average bus to traverse the route segment. Since χ buses per hour do so at C dollars per bus hour, total bus company costs are $CX/M/S$ and the cost per passenger served is CX/BS . The distance between stops is $1/Y$ miles. The maximum walk for any passenger is half this distance. Since origins and destinations are assumed to be uniformly distributed between stops, the average passenger would walk $1/4Y$ miles both to and from a stop or a total of $1/2Y$ miles. The cost of such a walk is $\alpha V/2Y$ dollars. The per passenger cost of time spent waiting at a stop is $\alpha V\beta/\chi$ dollars, while that of time in transit is MV/S dollars. Summing these four cost components gives the total cost per passenger for the steady state route:

$$[15] \quad Z = CX/BS + \alpha V/2Y + \alpha V\beta/\chi + MV/S$$

Suppose, for the moment, that overall bus speed, S , is independent of χ , the rate at which service is provided. Differentiating equation (15) with respect to χ , setting the result equal to zero, and rearranging terms would then yield

$$[16] \quad \chi = [\alpha V\beta SB/C]^{1/2}$$

as the cost minimizing value of χ : If speed were independent of level of service, the optimum service frequency would be proportional to the square root of the demand for service.⁷

In fact, if allowable stops per mile and passengers per mile-hour are held fixed, a reduction in the number of actual stops per mile

and hence an increase in realized speed would result from an increase in the number of buses per hour. That is, $\underline{x}$ and $\underline{S}$ are positively related. Thus, the optimal response to a doubling of $\underline{B}$ would be to increase $\underline{x}$ by a factor somewhat in excess of the square root of two.

Determining optimum service characteristics and hence minimum costs, then, requires specifying the relationship between realized speed, $\underline{S}$, on the one hand and, on the other, $\underline{x}$, $\underline{Y}$, $\underline{S}^*$ and the remaining parameters of the system. In each route mile, a total of $\underline{B}$ travelers per hour board and $\underline{B}$ leave $\underline{x}$ buses at $\underline{Y}$ or fewer stops. Hence, the average number of passengers that board or leave any one bus at any one stop is $\underline{\mu} = 2B/\underline{x}Y$. Suppose that people make travel decisions independently of each other. Then the probability that $\underline{r}$ people would board and alight at any one stop is given by the Poisson distribution with parameter $\underline{\mu}$. That is, $P[\underline{r}] = e^{-\underline{\mu}} \underline{\mu}^{\underline{r}} / \underline{r}!$. The probability that a given stop will be made, then, is one minus the probability that no one will be at that stop when the bus arrives, i.e., $1 - e^{-\underline{\mu}}$. The expected number of stops per mile is $\underline{Y}$ times this fraction. The expected time to travel one mile, $1/\underline{S}$, can therefore be written as the sum of the time actually absorbed in travel, $1/\underline{S}^*$, the time required to board and unload $2B/\underline{x}$ passengers and the time absorbed by the expected number of starting and stopping maneuvers:

$$[17] \quad 1/\underline{S} = 1/\underline{S}^* + 2B\varepsilon/\underline{x} + \delta Y [1 - e^{-\underline{\mu}}]$$

Equations (15) and (17) incorporate the normal bus company operating procedure of requiring travelers to walk to more or less

widely spaced bus stops. An alternative procedure that is sometimes followed allows passengers to hail a passing bus at the point along its route at which they may happen to encounter it. The prevalent procedure serves to reduce the number of stops a bus makes and hence to reduce both time in transit and the number of bus hours required to provide any specified number of bus trips. At the same time, however, this procedure requires passengers to incur walking costs. It seemed possible that, on lightly traveled routes, the frequency with which more than one passenger boards or alights at a given stop might be so small that the resulting savings in transit time and bus operating costs would not offset the loss in increased walking costs. On such routes, cost minimization would call for stops to be made on demand. To determine the circumstances under which this possibility might eventuate, the model summarized by equations (15) and (17) was altered to allow for an infinite number of possible stops. As $\underline{Y}$ approaches infinity, it can easily be shown that equations (15) and (17) respectively approach

$$[15'] \quad Z = C \chi / BS + \alpha V \beta / X + MV / S$$

and

$$[17'] \quad 1/S = 1/S^* + 2B(\epsilon + \delta) / \chi$$

The second cost model studied deals with a "feeder" bus route: Along each of the route's $\underline{M}$ miles, $\underline{B}$ people per hour board buses. All passengers dismount at the route's terminus, downtown. Using the same notation as that specified for the "steady state" route (except that

$\mu = B/\chi Y$), the average cost per passenger for the "feeder" route can be written:⁸

$$[18] \quad Z = C\chi/BS + \beta\alpha V/\chi + \alpha V/4\gamma Y + MV/2S$$

where:

$$[19] \quad M/S = M/S^* + 2MB\epsilon/\chi + \delta [1 + M(1 - e^{-\mu}) Y]$$

With an infinite number of allowable stops, these equations become

$$[18'] \quad Z = C\chi/BS + \beta\alpha V/\chi + MV/2S$$

and

$$[19'] \quad M/S = M/S^* + MB(2\epsilon + \delta)/\chi + \delta$$

Little would be gained by reproducing the derivatives with respect to χ and Y of any of these relationships. They are quite messy and do not yield explicit relationships for the cost minimizing values of χ and Y . It was therefore necessary to use iterative techniques to find these values.⁹ It is possible, however, to find an explicit short run marginal cost relationship once the cost minimizing values of χ and Y has been determined. If χ bus trips per hour are to be provided, $X = \chi M/S$ bus hours of service are required. Using this expression to eliminate χ in equations (15) and (17), multiplying (17) through by MB (total passengers per hour in the M mile route segment of interest), differentiating with respect to MB , and rearranging terms yields:

$$(20) \quad \begin{aligned} \partial(MBZ)/\partial(MB) = & \alpha V/2\gamma Y + \alpha V\beta M^2/SX + MV/S \\ & + 2VM^2B(\alpha\beta/X + 1) A_1 / [S(X - 2MBA_1)] \end{aligned}$$

where A_1 equals $\delta e^{-\mu} + \epsilon$ -- the time required to perform a stopping and starting maneuver times the probability that the stop at which the margi-

nal passenger boards would not otherwise have been made plus the time required to board him once the bus has stopped, $2A_1$ is the number of hours by which an additional passenger would increase the travel time of the MB/X travelers already aboard the bus he takes. Thus, $2MBVA_1/X$ is the cost he imposes on them. The first three terms on the right of (20) sum to the travel time costs of a trip -- costs which are borne by individual travelers. The fourth term, then, is the fare required to equate price and short run marginal cost. In addition to the cost a marginal passenger imposes on those already aboard the bus he takes, it includes the costs he imposes on all other travelers by reducing bus speed and hence the number of bus trips that can be provided by X bus hours. (18) and (10) yield an average feeder route fare of:

$$(21) \quad F = MNV(2\alpha\beta/X + 1)A_2/[2S(X - NA_2)]$$

where A_2 equals $\frac{2\epsilon + \delta e^{-\mu}}{2}$, an expression analogous to $2A_1$. Equation (21) is the exact fare only at the midpoint of the feeder route, a point at which there are $MB/2X$ passengers aboard the bus. Since no one is aboard a bus when it leaves its outer terminal, the fare there should be $MBVA_2/2X$ less than that given by equation (21). On the other hand, to set price equal to short run marginal cost would require the last person who boards a bus before it reaches its central business district terminal to pay $MBVA_2/2X$ more than the amount given by equation (21). That is, optimization of this sort of bus route would require fares to be inversely related to length of trip -- a strongly counter intuitive notion. Figure 4 shows the magnitude of this difference between feeder

route fares at the routes' inner and outer terminals.

In addition to the routes described above on which travel in one direction equaled that in the other, solutions were also obtained for routes on which five times as many trips were made in the main direction as in the back haul direction. The specific parameter values used are as follows. The data leading to selection of these values are described in the appendix.

- M, (Average) trip length: 3 miles
- Y, Walking speed: 3 miles/hour
- β, Wait for service as fraction of bus headway: 0.5
- V, Value of time in transit: \$1/hour
- αV, Value of walking and waiting time: \$3/hour
- S*, Bus speed when not stopping or starting: 20 miles/hour
- ε, Time to board or unload passenger: 1.8 seconds
- δ, Time to stop and start bus at bus stop: 18 seconds
- C, Cost of a bus hour's services: \$12.75 during morning
and afternoon peak; \$5.60 at other times.

The results of these computations are summarized in Figures 2-5 and Tables 1 and 2. Regardless of the specific combination of parameter values studied, both the steady state and feeder route models reveal considerable scale economies. Operating costs and flow rates that currently prevail during peak hours in the Twin Cities mostly underlie Figures 2 and 3. For a two direction average of 150 passengers per mile-hour (250 and 50 in the main and back haul directions, respectively -- roughly five times the current peak period average in the area), the long run marginal costs of main and

Cents Per
160 Trip

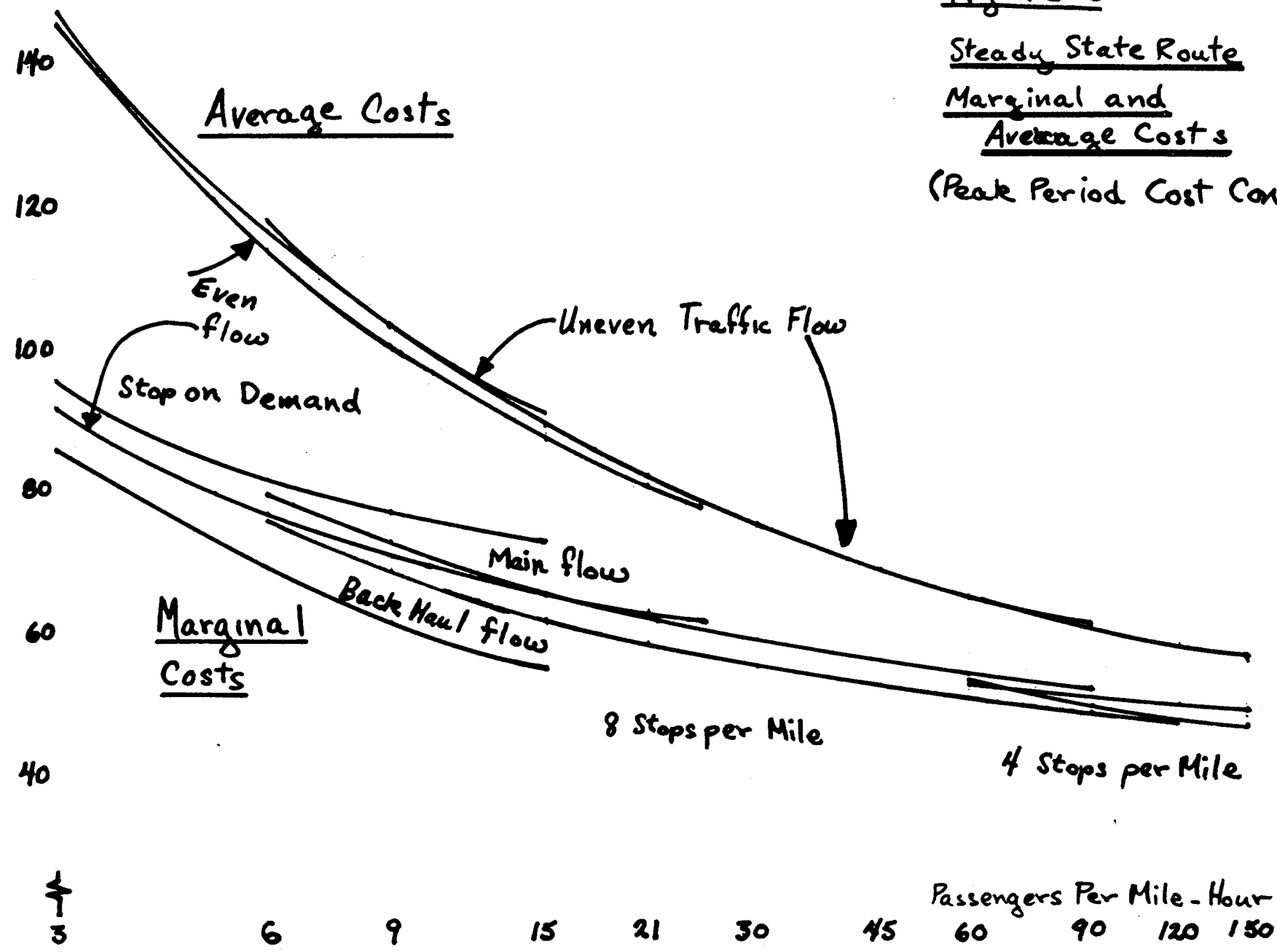
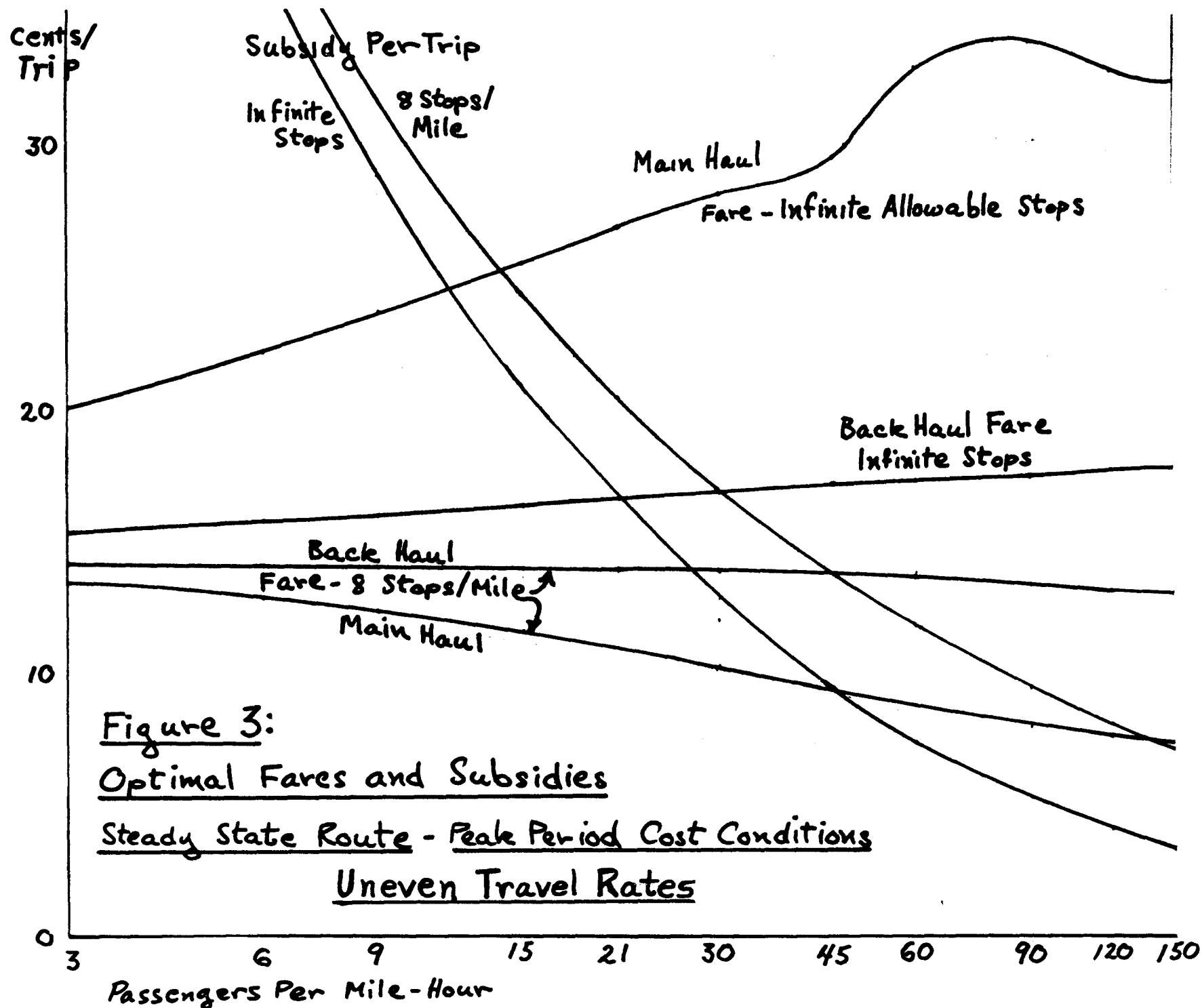


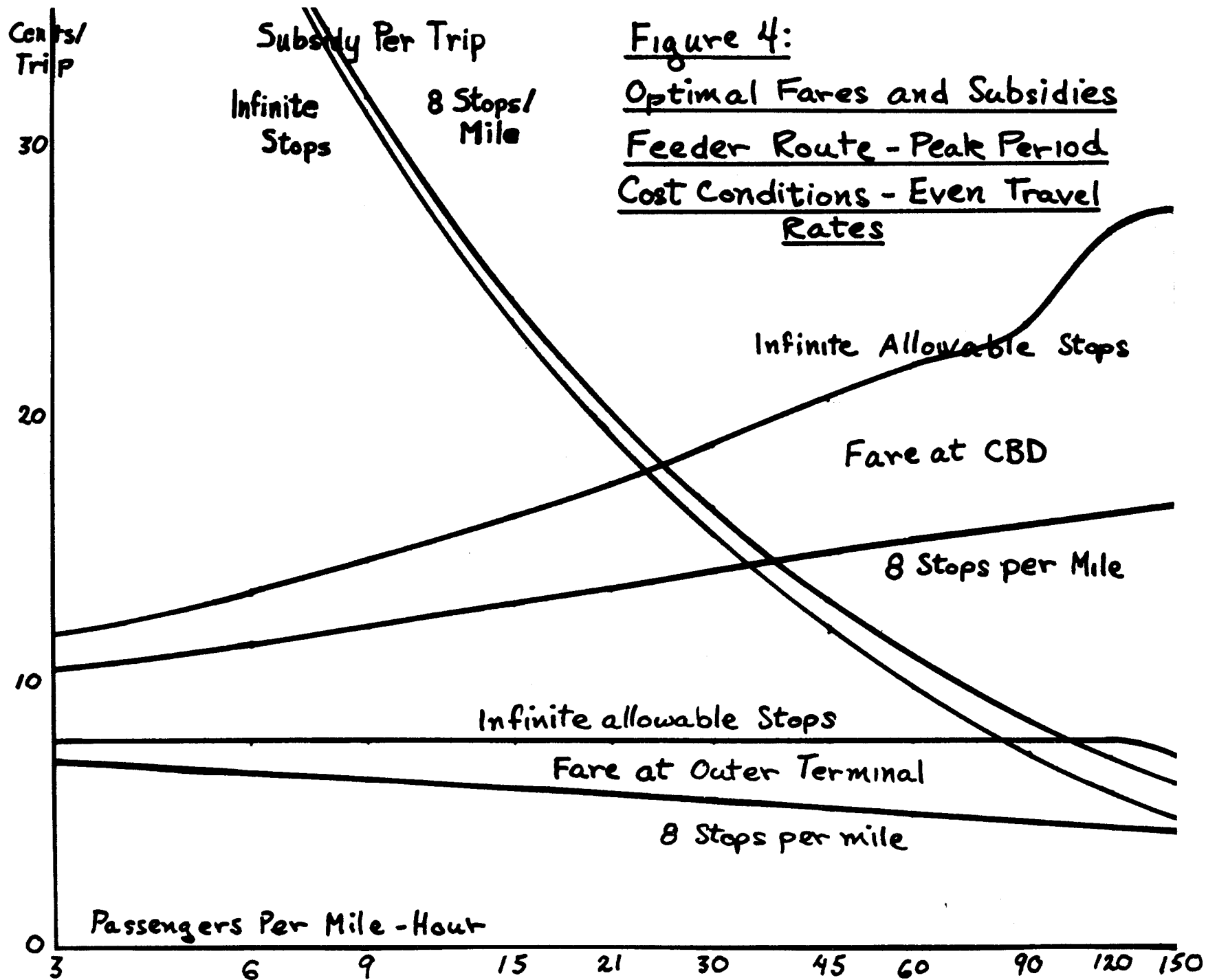
Figure 2
Steady State Route
Marginal and
Average Costs
(Peak Period Cost Conditions)

back haul trips are respectively 14 and 19 percent less than over-all long run average costs. Perhaps more important, even for this relatively high output rate, the weighted (by number of trips in each direction) average gap between long run marginal and average costs amounts to 57 percent of total bus company operating costs. At current travel rates -- about 30 and 10 passengers per mile-hour during peak and off peak periods respectively -- this gap equals 60-61 percent of optimal bus costs.

The differences between the long run marginal costs of main and back haul trips are surprisingly small as are the differences between the long run average costs of trips for even and uneven flow conditions. Regarding these latter differences, average costs for even and uneven flows were so close that it was impossible to plot both in Figure 2 except in that range of outputs for which total costs were minimized with an infinite number of allowable stops. This finding is related to the paradoxical results depicted at the bottom of Figure 3: For eight (or fewer) allowable stops per mile, the marginal cost fare for a trip is greater in the back haul than in the main haul direction.

The explanation for these curious results appears to be as follows: The number of passengers assumed to be aboard a back haul bus is a fifth that assumed for a main haul bus. This being the case, an additional traveler has a smaller effect on the time costs of these already aboard the bus he takes in the back haul than in the main haul





direction. But this "own bus effect" is not the only consequence of an additional trip. Since so few passengers per mile board the average back haul bus, the probability that adding a passenger will require an additional stop to be made is much greater for the back than for the main haul. Under the assumed conditions, it takes eleven times as long (19.8 as opposed to 1.8 seconds) to board a passenger if a special stop must be made for him than it would if the stop would have been made in the absence of his trip. By reducing operating speed and hence the number of trips that can be made with X bus hours per hour, additional stops affect all travelers, not just those aboard the bus in question. It would appear that, when the number of stops made per passenger boarded is appreciably less than one, this "system effect" of an additional back haul trip more than offsets its lower own bus effect.

Going through the arithmetic of a specific example may be worthwhile in this connection. To repeat, during the morning and afternoon peaks, approximately 50 and 10 passengers per mile-hour respectively board main and back haul buses on the average Twin Cities route. If stops are spaced an eighth of a mile apart, the cost minimizing service frequency for this output level is 8.88 buses per hour. With this service level, 16.9 and 3.4 passengers respectively would be aboard the average main and back haul bus. The own bus effect of an additional main haul trip accounts for 5.8 of the 10.2 cent marginal cost fare; the system effect for the remaining 4.4 cents. The corresponding

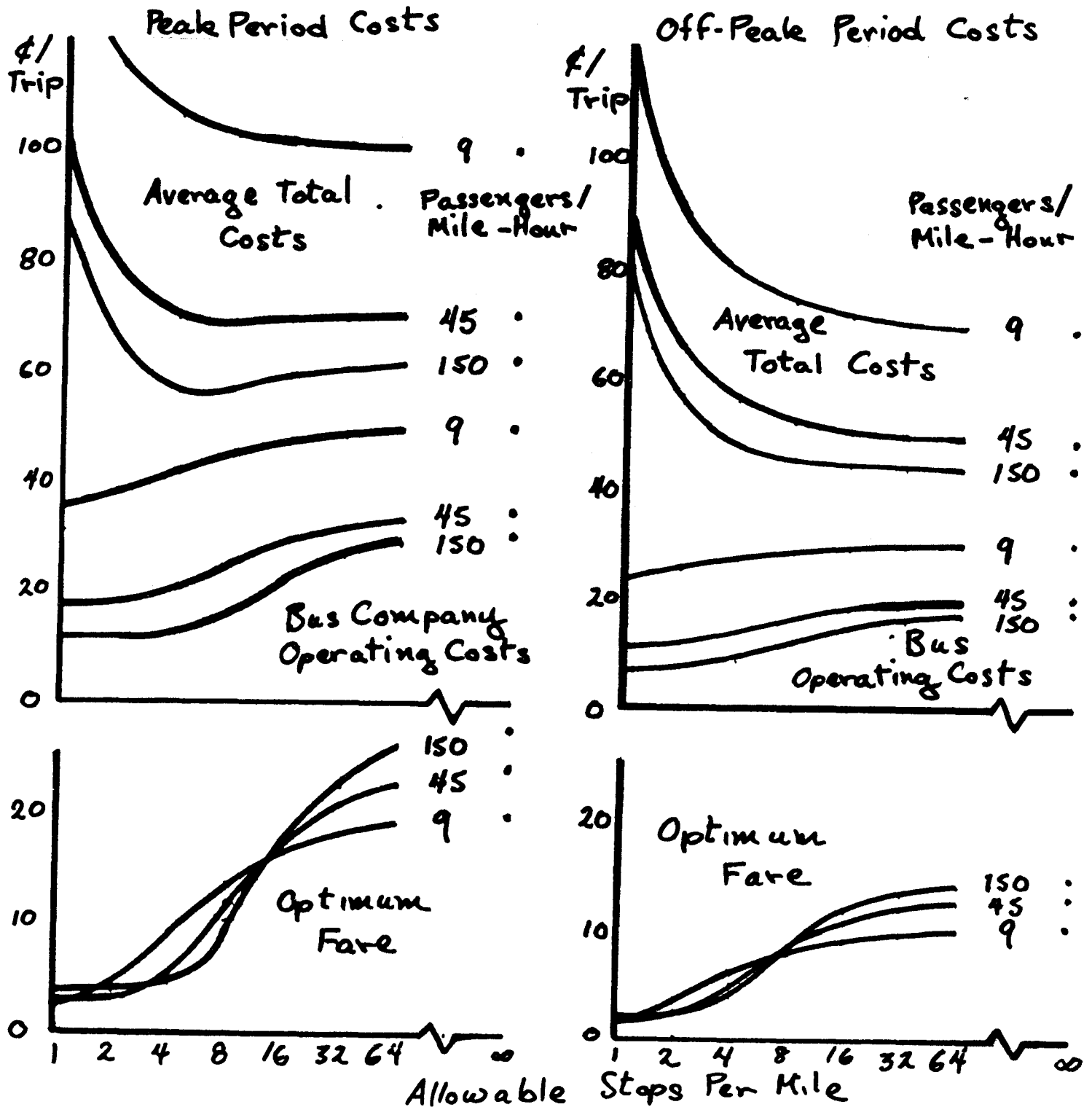
back haul figures are fare: 13.8 cents, own bus effect: 2.9 cents, and system effect: 10.9 cents. The substantial difference between the two system effects reflect the fact that the probability that boarding an additional passenger will require an additional stop -- $e^{-\mu}$ -- is 0.755 for the back haul but only 0.245 for the main haul.¹⁰

A point implicit in the foregoing should be made explicit: Stop spacing is a far more important determinant of optimal fares than is the rate at which trips are taken. Thus, as Figure 5 indicates, under peak load cost-even travel rate conditions, marginal cost fares for the steady state route only vary between 2.2 and 3.5 cents over the range 9-150 passengers per mile-hour when one stop per mile is allowed. At the other extreme, fares vary between 19.9 and 26.7 cents over this range of outputs when stops are made on demand.

Although not as dramatic as with fares, stop spacing has a substantial effect on optimal bus headways and hence on the bus operating and travel time components of total costs. These effects are particularly great for high trip output rates. Thus, under peak load cost-even flow conditions, the optimal service level for the steady state route with 150 passengers per mile-hour is 20.4 buses per hour when one stop per mile is allowed, 28.7 for 16 stops per mile, and 39.0 for an infinite number of stops. Bus company operating costs are 10.8, 22.9, and 30.6 cents per passenger under these alternative service rates, and the respective times costs for a three mile trip are 76, 36, and 35 cents.

Figure 5:

Average Total and Bus Operating Costs and
Marginal Cost Fares at Alternative Stop Spacings
(Steady State Route - Even Travel Rates)



With few passengers per mile-hour, $\underline{B}$, the sum of travel time and bus operating costs is a minimum if buses stopped when hailed. For large $\underline{B}$ values, between 4-8 stops per mile are optimal. Finally, for a narrow range of intermediate values, 16, 32, or 64 allowable stops would minimize total costs. Where the dividing lines between "small", "intermediate", and "large" values of $\underline{B}$ are drawn depends, of course, on the specific values given system parameters. As the travel time value used increases, the dividing lines occur at smaller and smaller $\underline{B}$ values. On the other hand, increases in trip lengths and in bus operating costs serve to increase these dividing lines. Thus, for off peak cost conditions, stopping on demand proved optimal for the steady state route even for the largest value of $\underline{B}$ tested, 250 passengers per mile-hour. However, for peak period cost conditions, 4-8 stops per mile provide minimum costs with 45 or more passengers per mile-hour. With $\underline{B}$ equal to 30 -- the present peak hour average in the Twin Cities -- 16 stops per mile is optimum, although average total costs for 8, 16, 32, 64, and infinite allowable stops all lie between 74.7 and 75.0 cents.

Table 1 provides data on optimum service frequencies for alternative bus cost, demand, and stop spacing conditions. In addition, denoting alternative $\underline{B}$ values by $\underline{B}_1$ and $\underline{B}_2$ and the associated optimal service frequencies by $\underline{X}_1$ and $\underline{X}_2$, this table gives the value of $\underline{E}$ which satisfies

$$(\underline{B}_1 / \underline{B}_2)^{\underline{E}} = \underline{X}_1 / \underline{X}_2$$

Table 1: Optimum Service Levels and Implied Scale

Economy Coefficients

Passengers/ Hour - Mile	<u>Peak Period Cost Conditions</u>				<u>Off Peak Costs</u>	
	<u>Unbalanced Flows</u>		<u>Balanced Flows</u>			
	<u>Stops / Mile</u>		<u>Stops / Mile</u>		<u>Stops / Mile</u>	
	8	∞	8	∞	8	∞
<u>Steady State Route</u>						
150	21.3	45.1	21.2	39.0	41.3	56.6
E	0.57	0.89	0.54	0.86	0.68	0.78
90	15.9	28.6	16.1	25.1	29.1	37.9
E	0.53	0.81	0.54	0.77	0.63	0.77
30	8.9	11.7	8.9	10.8	14.6	16.3
E	0.52	0.67	0.53	0.63	0.57	0.63
9	4.8	5.2	4.7	5.0	7.3	7.6
<u>Feeder Route</u>						
150	25.2	35.4	24.7	31.2	40.9	47.0
E	0.65	0.83	0.66	0.79	0.71	0.79
90	18.0	23.1	17.7	20.8	28.4	31.4
E	0.61	0.74	0.60	0.69	0.64	0.69
30	9.3	10.3	9.1	9.7	14.1	14.7
E	0.55	0.61	0.55	0.59	0.56	0.59
9	4.8	4.9	4.7	4.8	7.2	7.3

To repeat, equation (16) indicates that, if the rate at which travelers board a bus has no effect on the speed at which it operates, optimizing service frequencies would require $\underline{E}$ to equal 0.5. If steady state route-peak period costs-8 stops per mile conditions approximate those on a bus system, Table 1 indicates this "square root principle" to be a quite reasonable rule of thumb although an "0.55 principle" would be more nearly accurate. Other combinations of cost and route characteristics yield $\underline{E}$ values considerably larger than 0.5. With stop on demand operating rules, a separate stop is made for each passenger regardless of the travel rate. Therefore, for infinite allowable stops, optimum service frequency comes close to being proportional to the demand for service and scale economies approach (from above) those associated with reductions in waiting time that were discussed at the beginning of this section.

Table 2 constitutes an attempt to compare system attributes that currently prevail in the Twin Cities area with those that the steady state and feeder route models indicate would be optimal under similar demand and cost conditions. Going from current to optimal operating characteristics would call for the following: a) reducing peak period bus headways about a third, b) reducing off peak headways by about a half, c) reducing the current 30-cent fare by about 40 percent during the peak period and about 75 percent during the off peak period.

Table 2: Comparison of Optimum with "Current" Service Characteristics

	<u>Peak Period</u>				<u>Off Peak Period</u>	
	<u>Main Haul</u>		<u>Back Haul</u>			
	<u>Allowable Stops Per Mile</u>					
	16	∞	16	∞	16	∞
	<u>Steady State Route</u>					
Headway (min.)	6.0 min.	5.1 min.	6.0 min.	5.1 min.	8.0 min.	7.9 min.
Optimum Fare	17.6¢	28.1¢	15.5¢	16.8¢	8.8¢	10.1¢
Subsidy/Trip	15.1¢	12.8¢	15.1¢	12.8¢	20.1¢	19.7¢
Price/Trip	63.4¢	70.0¢	51.9¢	47.4¢	50.8¢	48.7¢
	<u>Feeder Route</u>					
Headway	6.2 min.	5.8	6.2 min.	5.8	8.3 min.	8.3 min.
Midpoint Fare	13.0¢	16.4¢	9.0¢	9.4¢	5.3¢	5.6¢
Subsidy	15.5¢	14.6¢	15.5¢	14.6¢	20.8¢	20.7¢
Midpoint Fare	53.4¢	55.0¢	43.1¢	43.4¢	37.2¢	36.2¢
	<u>"Current" Conditions</u>					
Headway	9.3 min.		9.3 min.		15.6 min.	
Fare	30.0¢		30.0¢		30.0¢	
Subsidy	?		?		?	
Price						
Steady State	79.3¢		66.8¢		74.8¢	
Feeder	68.9¢		60.5¢		67.4¢	

Putting these changes into effect would require increasing the present 600 bus fleet to about 850 and operating about 70 percent of these buses during the off peak period rather than about 35 percent as is presently the case.¹¹ If no increase in patronage were to result from these changes, the bus system would generate revenues sufficient to cover about 40 percent of its total costs. However, these changes would reduce trip prices considerably -- by about 20 and 30 percent respectively in the peak period main and backhaul directions and by 35- 45 percent during the off peak period.¹² Patronage would therefore almost certainly increase thereby making further service increases desirable and reducing the system deficit somewhat, at least on a per passenger basis.

To summarize, a subsidy to urban mass transportation systems reflecting the difference between average and marginal costs would probably not eliminate the decline in mass transit usage that has been experienced in virtually all urban areas. After all, currently used mass transit technologies provide services that are inferior goods to most income groups. As real incomes increase, the demand for these services will undoubtedly continue to decline. Still, a welfare maximizing subsidy policy would undoubtedly slow this movement and might even hasten the adoption of new technologies that promise vastly improved service characteristics.

Appendix: Derivation of Cost and Related Parameters

Explanations are in order for some of the parameter values used in optimizing bus route service characteristics and in describing current operating characteristics.

A variety of studies [see, e. g., Beesley, Gronau, and Lisco] have concluded that the amounts travelers appear willing to pay to save time aboard mass transit and other vehicles vary with their wage rates. The fraction varies from about 25 percent for low income travelers to about 50 percent for middle and upper income groups. In turn, people appear willing to pay between 2-3 times as much to save walking and waiting time as to save time aboard conveyances. [See, e. g., Lisco, pp. 79-88]. Bus travelers largely come from low income groups. The \$1 and \$3 an hour used for time in transit and walking and waiting time respectively therefore seem reasonable albeit perhaps a bit on the high side.

\$12.75 and \$5.60 are approximately the marginal costs to the Twin Cities Metropolitan Transit Commission of standard 55 passenger bus hours during peak and off peak periods. These estimates were determined as follows: A new 55-passenger bus costs approximately \$37,000 and is depreciated over a 12-year period. Applying a 10% interest charge to $\$37,000/2$ and adding depreciation of $\$37,000/12$ yields an annual capital cost of \$4933. Dividing by 1560 peak hours a year -- 6 peak hours per weekday times 5 weekdays per week times 52 weeks per year -- yields a peak hour capital cost of \$3.16.

Bus driver costs account for about 70% of total system costs. Taking into account fringes, overtime, etc., the average cost of a driver hour is currently about \$6. Adding a peak bus driver hour involves overtime and other premia that would not be paid for an off peak driver hour. The Transit Commission's labor contract is extremely complex. It involves, inter alia, guaranteed minima of 6 hour days and 40 hour weeks and time and a half for both hours in excess of 8 per day and beyond 11 hours from starting time. It is therefore impossible to determine the exact marginal cost of a peak hour driver. Eight and four dollars for peak and off peak hours respectively seem as reasonable guesses as any. Allocating fuel, tire, maintenance, administrative, and overhead costs on a per bus hour basis adds \$1.58 to the above figures.

Twin City Lines (the predecessor of the Metropolitan Transit Commission) bus trips had mean and median lengths of about 3.5 and 2.3 miles respectively in 1962. For the system as a whole during the average weekday, B, passengers per mile-hour, appeared to be approximately 11 between 9 a.m. - 3 p.m. and about 30 between 3 - 6 p.m. (Derived from Minnesota Department of Highways, pp. 25, 40.) Back haul travel during the peak hour appears to take place at the level characteristic of off peak hours.

In both London and New York, the time required to decelerate, open doors at a stop, and accelerate, δ, averages approximately 21 seconds. ε is approximately 1.5 seconds for unloading in New York and for both loading and unloading in London. In New York, the time absorbed

by fare collection results in a 2.6 second value for $\underline{e}$ in loading.¹³

The estimates of current service characteristics employed in Table 2 and in inferring the magnitude of scale economies under a "service proportional to demand" policy were determined as follows: A sample of points was drawn from the 28 traffic analysis districts covering Minneapolis and its immediately adjacent suburbs that were employed in the 1958 Twin Cities Area Transportation Study. Almost all (95 percent) of the bus trips taken during the survey period had "home" as either origin or destination. The number of observations drawn from each district and the weights attached to each observation were therefore based on the number of trips originating at home that did not have "school" as a destination. School destination trips were eliminated because the majority of them are taken in other than mass transit buses.

Approximately 65 percent of all non-school trips taken in the survey area had an origin or destination in either the St. Paul or Minneapolis central business district and most bus routes either terminate in or go through one of the CBD's. Peak and off peak service frequency estimates for each sample point were therefore based on the number of buses scheduled toward the central business district during the period 6-8 a.m. and 8 a.m.- 4 p.m. respectively on the nearest street having bus service. In the few cases where the nearest bus route did not provide service to the Minneapolis CBD, the buses counted were those heading away from the nearest terminal. See Minnesota Department of Highways, pp. 4, 9 - 14, 73 - 74.

FOOTNOTES

* Professor of Economics, University of Minnesota and York University.

I am indebted to Lynn Gerber, Director of Scheduling, Twin City Lines, and to John Jamieson, Director of Transit Development of the Twin Cities Metropolitan Transit Commission for providing data, to Marvin Kraus for very helpful comments, to Myrna Wooders for computer programming and checking algebra and, for financial support, to a grant from the Departments of Transportation and Housing and Urban Development to the University of Minnesota Center for Urban and Regional Affairs.

1. Some important second best justifications for mass transit subsidies also exist. These are not dealt with here.
2. The first part of this section is a modest modification of Robert Strotz' "First Parable."
3. Implicit in this formulation is the assumption that the cost of a bus hour is independent of its carrying capacity. While clearly unrealistic, this assumption is not as bad as may at first blush seem to be the case. An increase in the capacity of a bus results in a less than proportionate increase in its capital, fuel, and related costs. More important, driver wages and fringe benefits account for about 70 percent of the total costs of a typical urban bus operation. The cost of a driver hour is independent of the size of the bus he operates.
4. Excluding trips to school and home, just over 50 percent of all bus trips in the Twin Cities area had "work" as a destination in 1958, while an additional 25 percent involved social-recreational or personal

business purposes (see Minnesota Department of Highways, p. 26). Many trips in these three categories require being at the destination point at a specified time. To the extent that bus schedules do not fit in with such time constraints, a wait at destination is necessary.

5. If the effect additional bus travel might have on highway congestion and hence on travel time can safely be ignored.
6. These data and the numbers which follow are derived from Minnesota Department of Highways, pp. 4, 9 - 14, 40, and 73 - 4. See the Appendix for a discussion of how they are developed.
7. William Vickrey first propounded this square root principle to me. I presume that he based his assertion on a similar analysis. Actually, Section II of this paper turns out to be merely an elaboration on the middle paragraph of his 1955 article.
8. For such a route, it is quite likely that optimum stop spacings and service frequencies would vary with distance from downtown. These possibilities are ignored in what follows.
9. For reasons suggested by Figure 5, it was impossible to solve simultaneously for the optimizing values of $\underline{X}$ and $\underline{Y}$. The procedure finally settled upon was that of using Newton's method to determine the optimum service frequency for each of a variety of stop spacings and trip output levels. The long run marginal costs of providing $\underline{B}$ trips an hour was approximated by finding the cost minimizing values of $\underline{X}$ associated with $\underline{B}$ and $1.05 \underline{B}$ and then dividing the difference between the two cost levels by $.05 \underline{B}$. Costs determined in this fashion typically differed from short run marginal time costs under optimal conditions by less than half a mil.

10. It is perhaps worth noting that, if 16 rather than 8 stops per mile are allowed under the conditions dealt with in this paragraph, the paradox would disappear: Marginal cost fares would then be 17.6 and 15.5 cents respectively in the main and back haul directions.
11. For reasons suggested by the Appendix discussion of the derivation of the \$12.75 and \$5.60 peak and off peak period costs of a bus hour, an increase in the off peak fleet utilization rate from 35 to 70 percent would result in both a reduction in the cost of a peak period bus hour and an increase in the cost of an off peak hour. These changes, in turn, would make optimal somewhat better peak hour service and somewhat poorer off peak service than levels indicated in Table 2.
12. The alternative estimates of current trip prices equal the current fare for a trip plus the time costs that would result from the steady state and feeder route models given current travel rates, 9.3 and 15.6 minute headways, and 16 allowable stops per bus mile.
13. I am indebted to T. M. Coburn of the United Kingdom Road Research Laboratory for these data and also for noting that two-door operation characterizes most American bus services. As a result, total loading and unloading time at a stop is approximately equal to the greater of 2.6 seconds times the number of boarding passengers and 1.5 seconds times the number of dismounting passengers. This complicating factor is ignored in what follows.

REFERENCES

1. William Baumol, "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," American Economic Review, June 1967, 57, 415-426.
2. Michael Beesley, "The Value of Time Spent in Traveling: Some New Evidence," Economica, May 1965, NS 32, 174-185.
3. Reuben Gronau, The Value of Time in Passenger Transportation: The Demand for Air Travel, New York, 1970.
4. Thomas Lisco, The Value of Commuters' Travel Time: A Study in Urban Transportation, Unpublished Doctoral Dissertation, University of Chicago, 1967.
5. William Vickrey, "Some Implications of a Marginal Cost Pricing for Public Utilities," American Economic Review, May 1955, 45, 605-620.
6. Minnesota Department of Highways, The Role of Mass Transit: Twin Cities Metropolitan Area, St. Paul, 1963.
7. Robert Strotz, "Urban Transportation Parables," in Julius Margolis, ed., The Public Economy of Urban Communities, Washington, D.C., 1965, pp. 127-169.