REDUNDANCY IN PUBLIC TRANSIT

Vol. II: The Profits of Competition in Public Transit

Philip A. Viton

Monograph 029

August 1980
Final Report

Institute of Urban and Regional Development
University of California, Berkeley

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
Urban Mass Transportation Administration
Washington D.C. 20590

REDUNDANCY IN PUBLIC TRANSIT

Vol. I	On the Idea of an Integrated Transit System	Martin Landau Donald Chisholm Melvin M. Webber	MG-030
Vol. II	The Profits of Competition in Public Transit	Philip A. Viton	MG-029
Vol. III	The Political Economy of Transit in the San Francisco Bay Area, 1945-63	Seymour Adler	MG-030
Vol. IV	Structure, Competition, and Reliability in Planning and Operations	Jonathan B. Bendor	MG-032

REDUNDANCY IN PUBLIC TRANSIT

Vol. II: The Profits of Competition in Public Transit

Philip A. Viton
University of Pennsylvania

Monograph 029

August 1980
Final Report

Institute of Urban and Regional Development
University of California, Berkeley

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
Urban Mass Transportation Administration
Washington D.C. 20590

The study was supported by the
U.S. Department of Transportation,
Urban Mass Transportation Administration,
Grant No. CA-11-0001.

Additional support was provided by the

UPS Foundation

Center for Transportation Policy Research
Institute of Transportation Studies

and

Institute of Urban and Regional Development

TABLE OF CONTENTS

Summary

BART has already been built. No matter what operations policy is followed, how many riders the system attracts, or how cost-effectively it is run, the capital costs have been incurred, and the bonds and other instruments issued to obtain capital funds will have to be paid. From the point of view of day-to-day systems operations, these capital costs may be treated as irrelevant.

This report asks a series of questions about the operations policy of a rapid rail system. Specifically:

- It was promised that BART would cover its variable (operating) costs from the farebox. It has not done so. Was this an unreasonable demand?
- BART now competes with a highly used bus transit system which is heavily subsidized. But when BART was conceived no such system existed. What can BART do when faced with this sort of competition?

- On these routes where BART and AC Transit compete directly, the levels of service are virtually identical. Does this have anything to do with the operations losses? And is there an operations plan which, by differentiating BART's from AC's level of service, would enable the original promise to be met?

- If BART and AC Transit were to merge and form a transit "super-agency", would there be an incentive to eliminate bus service entirely?

- What about automobile travel? The effect of BART in peak-period travel has been to make substantial inroads into bus travel volumes, but comparatively little effect on total auto travel is noted. Is this to be expected?

From the nature of these questions it is immediately clear that:

- They are "what-if" questions. That is, they ask about the possibilities open to the system, possibilities which are not actually observed.

- Their answers require some sort of representation (model) of Bay Area travel possibilities.

The bulk of the report sets up the model and solves it for a number of transit scenarios in order to answer the questions posed. To contribute to the answers, a model of transportation in the Bay Area should recognize that:

- Consumers have preferences over different modes and modal characteristics; and have a variety of options open to them. Any discussion which ignores consumer behavior is doomed to inadequacy.

- Transit operations are characterized by an essential interdependency. What BART does will not only determine its own patronage and revenues, but will also alter the patronage and revenues of competing modes. Where BART competes directly with a bus system, to assume that the bus agency will not alter its operations plan in response to BART decisions flies in the face of reasonable expectations.

- The questions as posed here assume a motivation on the part of the transit agencies which is not necessarily the principle now guiding their actions. This motivation is to focus on loss-minimizing behavior. That is, a particular type of "what if" question is asked here: what would happen if BART (and possibly, though not necessarily, AC Transit or any other competing bus agency) acted in such a way as to minimize, or even more than cover, operating losses? Focussing on operating costs is a consequence of the short-run irrelevance of capital costs pointed out above.

The report suggests a framework embodying these considerations with the following features:

- Consumers choose their preferred travel mode from among six combinations of access and linehaul modes. These are: auto alone; bus with auto access (kiss-and-ride); bus with walk access; BART with auto access; BART with walk access; and BART with feeder bus access.
- The BART District is conceived as operating its own feeder service, and thus providing what has been called "integrated" service. This allows us to envision the possibility that BART may be motivated by loss-minimizing concerns to operate a different type of feeder service than that offered by a bus agency with its own (competing) concerns.
- BART freely sets its operating characteristics to cover all variable (operating) costs. These characteristics are the fare on BART, the headway between successive trains, and possibly the train length as well. When it is in BART's interest to run a feeder service, it freely selects the bus fare, bus headway, and ease of feeder access.
- The competing bus system is conceived as having two possible behavioral motivations. The first is loss-minimizing behavior analogous to BART's, in which fare, headway, and access ease are freely selected.
- A system charged to minimize loss with respect to operating costs can cut operating losses to zero by not operating at all. Of course, capital costs would still be incurred; but these are irrelevant, since they are sunk costs.
- Zero output may be unacceptable as a public policy. Thus we also conceive of the bus agency as pursuing the least-cost way of achieving a specified minimal level of service at some maximum fare. This may require money-losing (i.e., subsidized) bus operations. We may thus discuss BART's possibilities in the face of a subsidized bus agency.

The analysis resulting from these considerations is set out in the bulk of the report; and the tables of Section 5 give some numerical answers to the "what-if" questions. We may summarize the major conclusions in the form of some tentative answers to some policy questions, as follows:

- BART was originally designed to cover operating costs from the farebox. In retrospect, was this unreasonable? Not at all. The results presented here, especially in Table 8, show that there is room for both BART and a competing bus system in the Bay Area transit market; and that cost-covering operations are in many cases feasible for both modes. Of course, this result is to some extent dependent on the travel characteristics of the corridor in question--the physical details of the corridor as well as the demographic characteristics of the travelers--but the basic result holds. BART can cover its costs from the farebox.

- Would this result still hold if BART had to compete with a heavily subsidized bus transit system? In many cases, yes. As Table 9 indicates, competition with a money-losing (hence, subsidized) bus system is possible.

- What changes would have to be made in BART's operations for this to occur? Basically, BART would have to differentiate its product from that of the bus system, something which, as Table 1 indicates, it has thus far not tried to do. This differentiation takes the form first, of a greatly improved access system, to circumvent the difficulties of gaining access to a fixed-route mode. Second, running shorter trains more frequently has two benefits: it reduces BART's excess capacity and it has the effect that average waiting times of users is lowered, thus making the system more attractive to potential riders. Third, judging from the available evidence on consumer preferences, riders are prepared to pay premiums for this service: the equilibrium (profit-maximizing) fares are considerably above current levels.

- Such high fares may simply be infeasible. Is there a way to keep fares at current levels, and adjust other level-of-service parameters to attract enough riders to cover costs? It appears that this is not a practicable option. A certain fare level seems an essential ingredient of cost-covering operations.

- Would not a transit super-agency, running both transit modes as a transit monopoly, find it optimal to close down one system altogether? No. As is apparent from Table 10, a transit super-agency, if it had the same loss-minimizing motivations as the individual carriers, would never find it optimal to close down one transit mode when, acting independently, both could cover costs.

- *But isn't it a waste of resources to have two transit systems service a single corridor?* There are three reasons why this is not the case. First, it appears that there is an optimal degree of product differentiation in the economy (Lancaster, 1975); and only in exceptional circumstances will a single product .be optimal. Note that the sort of loss-minimizing equilibria considered here result in transit "products" differentiated along both fare and service-characteristics dimensions, in line with the qualitative thrust of Lancaster's results. Thus, whether the relatively undifferentiated products now offered by BART and the competing express-bus services represent an efficient use of resources, is doubtful. Second, it may appear odd that the super-agency would *want* to maintain both systems. But a moment's reflection on the structure of the economy at large suggests that it is often good for a firm to produce competing products. To take a few obvious examples: General Motors is organized into divisions which directly compete with one another. The same holds for dry-cereal manufacturers. And recently, under recordings of Beethoven's Fifth Symphony, one company (Columbia) was represented with five different versions, two (Phillips and RCA) with three. So producing competing (and differentiable) products is by no means surprising. The third reason is one touched on in the introduction to the paper, namely, the economic value of redundancy. Since it appears that in many instances the sort of competition modeled here would lead to both systems being operational, this idea has only appeared indirectly. However, it remains a reason to consider seriously what appears on its face to be wasteful duplication.

- *Doesn't it contravene our notions of public policy to have differentiated (high-fare/high-quality and low-fare/lesser-quality) services as are suggested here?* This is to some extent a question of empirical politics. However, it should be noted that differentiated products in industries generally thought to be public utilities are by no means unknown. For example, the Postal Service now offers highly differentiated services (first class, registered, express, airmail delivery) and our notions of public policy do not appear to suffer violence. Again, electric and gas utilities often offer discounts in price in exchange for a lessening of service quality (for example: being first to be cut off the network in case of emergency, or agreeing not to place burdens on the network during the peak). In transportation, Amtrak offers differentiated passenger service: conventional versus Metroliner trains (differentiated largely on a travel-time dimension) and coach versus club-car service (differentiated on a dimension of amenities while riding) at differing fares. So the existing spectrum of public endeavors is by no means free of product differentiation of the sort suggested here. These do not appear to be objectionable. Whether a differentiated transit structure is politically viable in the Bay Area is of course another question. Still, there is no a priori reason not to consider it. And since it can result in losses on some systems being eliminated, such a structure might be all to the good.

1. Introduction

1.1 The Problem

It was clear from the start that the new BART system would face substantial competition from other modes. Chief among the competing modes was the private car. Bay Area auto-ownership rates are among the nation's highest, and reliance on the private car was fostered by an early full-freeway system.[1] With a distribution of incomes skewed towards the upper tail[2] there was every reason to expect that substantial numbers of commuters would elect to remain in their cars.

But BART did not face competition only from automobiles. In the East Bay, service provided by the Alameda-Contra Costa Transit District (AC Transit) is also competitive. AC Transit was the successor to the Key System, which provided local East Bay service as well as transbay service on rails over the San Francisco-Oakland Bay Bridge. In 1955 the Key System petitioned the California Public Utilities Commission to allow abandonment of the transbay lines, claiming they lost money.[3] In 1956 the AC Transit District was formed to coordinate transit in the East Bay. Two years later the Key System ran its last trains over the Bay Bridge, and on October 1, 1960, AC Transit bought the physical assets of the Key System and its parent company, the Bay Area Public Service Corporation.[4] AC Transit immediately began rennovating the old Key System buses, and adopted a policy of gradual replacement of those vehicles.

Since that time AC Transit has provided what has been widely regarded as extremely effective transit services. The system provides extensive coverage of the East Bay, with express buses running to the principal East Bay CBD in Oakland as well as both local and express service to San Francisco. The district has gained a reputation for cost efficiency, a reputation which is confirmed by comparing its operating costs to those of other bus transit properties.[5] In an era of increasing reliance on the private car, AC Transit managed to retain its ridership, and even increased patronage by about eight percent over the period 1962/1963 to 1976/1977.[6] Thus, BART could expect substantial competition from the bus system as well as from the car.

It is interesting to note in this regard that BART's service quality (including the fare) closely parallels the competing bus system's. Table 1 shows those service characteristics directly set by the management of the two systems, for a transbay trip from Berkeley to San Francisco, where there is direct competition between the two modes. As is apparent from that table, service characteristics are quite similar: BART is slightly faster and costs slightly more than regular bus service; for express buses, quality of bus service is virtually identical to that of BART. The headway comparisons must be read with caution. AC Transit runs scheduled service and is reputed to adhere strictly to its schedule. At present, BART headways are subject to considerable variability.

TABLE 1

Berkeley-San Francisco Travel, 1977[1]

	Headway	In Vehicle Travel Time	Fare
BART[2]	6 mins.	19 mins.	$ .65
Bus[3]--regular	10 mins.	23 mins.	.55
Bus[3]--express	9 mins.	19 mins.	.65

Notes

1. Peak travel only (data are for the morning peak). Travel between Berkeley and Montgomery Street.

2. BART headways as planned for full operations.

3. Travel between University Avenue/Berkeley and Transbay Terminal/San Francisco.

The service operated by AC Transit has failed to cover its operating costs in any year since its founding in 1960.[7] BART was expected to rely for its construction costs on general obligation bonds secured by the property tax. On the other hand, it was expected to pay for all rolling stock and operating expenses out of the farebox.[8] The question may be raised, whether this was a reasonable expectation, in light of the fact that BART provided a comparable level of service to that provided by a deficit bus system.

Recent developments in transit supply patterns give point to this doubt. In February of 1968, the Board of Estimate of New York City awarded the first franchise to a private bus company to provide express bus service between Fresh Meadows in Queens and midtown Manhattan.[9] The experiment was a success, with the result that there are now some fifty express bus routes linking Manhattan and the suburbs. Some are run by private companies, some by the Transit Authority (TA) of New York City, but all share three characteristics. First, they are all profitable. Second, the quality of service is high: all routes provide peak service at short headways, good accessibility, and high probability of getting a seat.[10] Third, the fare, though controlled by the Board of Estimate, is considerably higher than the normal New York City TA fare: recently the express bus fare was $1.50 while the fare for regular TA service was $.50 . Nor is this development limited to bus lines. Among the properties providing rail-based commuter service, the Illinois Central Gulf, the Chicago and Northwestern Railroad, and the Lindenwold Hi-Speed Line come close to meeting operating expenses out of farebox revenues.[11] These systems are also characterized by high levels of service and high fares. By contrast, in 1975-1976 BART had operating costs of $59.0 million, and gross fare and concession revenue of $21.7 million, thus covering only thirty-seven percent of its costs from revenues.[12]

BART has already been built. There is no immediate reason to re-examine the question of the total resource costs expended in providing trips by different modes.[13] Rather, the question now is how, in the short run, might BART adjust its variable factors of production to reduce (if not cover) its operating losses? It is important to realize that the Bay Area's transit industry has three main modes: the rail systems of BART and the Southern Pacific; bus systems run by AC Transit, Golden Gate Transit, SamTrans, Santa Clara County Transit, and San Francisco's Muni; and the private car. To assess the prospect of loss-minimizing service, one needs to simulate the interdependencies among the three sets of carriers. Further, it is important to know just how the users of the system will react to short run changes in service quality. What is needed, in short, is a full multi-modal demand and supply model of Bay Area transit.

Besides answering the initial question, whether it is possible for BART to live up to its promise to cover operating costs from the farebox, such a framework would provide a way of answering some important organizational questions. First among these is the relation between BART and AC Transit in an era of fiscal limits. As already indicated, neither system at present covers its operating costs. As shown in Table 2, BART's 1975-1976 loss per passenger was $1.13 , compared to $0.48 on AC Transit. From this it may seem that of the two agencies, the financial situation of BART is the more precarious. But this is wrong.

In June of 1978 the people of California approved Proposition 13. This measure, placed on the ballot by initiative, imposed strict limitations on the rates of property tax collected by local governments. While in the fiscal years 1978/1979 and 1979/1980 state budget surpluses cushioned the effect of reduced property tax revenues, it is unlikely that this situation will persist. The eventual effects of the property tax reduction will be different for BART and AC Transit: Table 2 shows why. In 1975/1976 BART received over half its nonoperating revenue from the use tax,[14] and only thirteen percent from property tax monies. By contrast, AC Transit obtains sixty percent of its nonfare funds from property tax levies,[15] with the remainder split between government grants and an apportionment of local transportation funds generated within Alameda and Contra Costa Counties.[16] Thus, the effects of reductions in the property tax are likely to be much more severe at AC Transit than at BART.

TABLE 2

Operating Characteristics

1975/1976

REVENUES, COSTS, AND LOSSES	BARTD[3]	AC Transit[2]
Fare & Concession Revenue[1]	$21.7 million	$18.4 million
Operating Cost	59.0 million	46.4 million
Loss	37.3 million	28.0 million
Loss Per Revenue Passenger	$1.23	$ .48

SOURCES OF FUNDS TO MEET LOSS		
Property Tax	$ 5.0 million (13%)	$19.4 million (60%)
Transactions and Use Taxes	21.1 million (57%)	-
Governmental Aid	1.6 million (4%)	5.9 million (18%)
Other[4]	9.6 million (26%)	6.8 million (22%)
TOTAL:	$37.3 million (100%)	$32.1 million (100%)

NOTES:

1. Including contract service (13% of revenue) for AC Transit.

2. Source: AC Transit Annual Report.

3. Source: BARTD Annual Report, 1976.

4. Net of bond principal and interest. Includes, for BART, interest (including "construction funds, interest and other,", $2.9 million), $4.5 million; borrowing on capital account, $5.1 million. For AC Transit: apportionment of local transportation funds.

Already by the summer of 1978 both agencies were aware of the difficulty. Some of the ideas mentioned informally involved some sort of merger of the two systems--perhaps under the auspices of the Metropolitan Transportation Commission. Thus, it was thought, AC Transit might reap the benefits of access to funds from sources other than the property tax. The question arises as to the loss-minimizing operating characteristics of such a combined system: might it not entail shutting down one or the other system entirely? In a companion volume, Landau, Chisholm, and Webber point to advantages of redundancy in systems--that when one component is out of commission a hitherto redundant component may be able to continue providing service. This consideration is particularly relevant in the case of Bay Area public transportation, where both BART and AC Transit are periodically shut down for extended periods because of labor disputes. If it turns out that the action which minimizes short-run losses for a combined system is to rely on one component alone, then it is worth considering the social benefits of retaining redundancy. But first we must see if the one-mode outcome might be preferred.

A second organizational question involves coordination between the two systems. BART, as Webber (1976) has noted, "offers one route in each compass direction, and hence only limited distribution across the urbanized area it serves;"[17] and in designing the system with widely-spaced stations, the planners sacrificed ease of access to high vehicle speed. It was clear that to function well, BART would have somehow to provide--either on its own or in conjunction with another transit operating company--easy access to the trains. The way chosen was to link the BART system with whatever bus operation was locally available--AC Transit in the East Bay and the Municipal Railway (Muni) in San Francisco. The degree of coordination was always a matter of some dispute; as Adler (1978) has noted "all the public statements made during the . . . 1950s and through the passage of the BARTD bond issue in November 1962 would affirm (the) idea of complementarity. . . the regional system would be dependent on the delivery of patronage through feeder and local service transit operations."[18] On the other hand, no arrangements for coordination were written into the enabling legislation of either AC Transit or BART. When BART opened, AC Transit did redirect some of its buses to bring them close to the BART stations; but there has been constant dispute between the two systems as to the degree to which buses should serve BART stations. At present, there is a free transer from BART to AC Transit, but not in the other direction.

There are, of course, technological difficulties in implementing reduced-fare transfers from general-route buses to BART. Nonetheless, the question arises whether BART could do better given a better transfer system. That AC Transit might not find it in its own best interest to provide such a service is clear; on many East Bay routes the two systems compete directly at comparable levels of service. Given the relatively low accessibility of BART, the only real advantage of the bus system on routes where direct competition takes place lies in its ease of access. So it is worth asking what kind of feeder service (and at what fare) would do best for BART.

1.2 Scope and Methods

In the previous section I identified three broad questions of current interest. First is the question of the self-sufficiency of Bay Area transit modes: to what extent are they able to cover their operating costs, and how should service characteristics be set to do so. Related to this is the question whether BART is wise to offer service very little differentiated from that of its principal transit competitor, the bus. The second is an organizational question, predicated on the declining availability of property-tax funds to support bus-system deficits. What would be the consequences of merging the two agencies, and what sorts of service could we expect from a transit monopoly? The third question involves the best level of provision of a feeder service to the rapid-rail system. In this section I shall discuss in broad outline a way of approaching these questions.

Two considerations motivate the analysis. The first of these is the notion of competition. For reasons stated above, it is important that we model consumers as having a wide choice of transportation modes available to them. In the anlysis which follows I allow for a total of six combinations of linehaul and access modes: auto only, bus with walk access, bus with auto access, BART with walk access, BART with bus access, and BART with auto access. It is also important to focus on travel in which all these options may reasonably be expected to obtain. For this reason I place the analysis in the context of a corridor route, in this case a route connecting a residential area with a CBD or major shopping center. There are several Bay Area transportation corridors in which the auto, bus, and BART modes compete; among them the Berkeley-San Francisco, the Richmond-Oakland, and the Richmond-San Francisco corridors. In each of these a bus route and a major freeway closely parallel the BART line. On BART's Concord line, the major source of competition is the private car, since Greyhound bus service has been severely curtailed: I shall not discuss the more restricted case of BART-auto competition. For an analysis of the possibilities of bus-auto competition see Viton (1980).

The second consideration which will play an important part in the analysis is the notion of loss-minimization. This concept is the driving force in the design of systemic service levels. It may be approached in two ways. The first, absolute loss-minimization, implies that unless a negative loss (excess of total revenue over total cost) is achieved, service should cease. Since zero service can be achieved at zero (variable) cost, the no-loss output is a lower bound. However, it is clear that many urban transportation services are constrained to provide at least some minimum level of service--say at least five buses per hour at some maximum fare. In these cases, one wants to investigate the constrained loss-minimizing service characteristics, subject to given minimal service. To anticipate a later result, loss-minimizing transit fares may be well above currently observed fares. If these fares are considered infeasible (for whatever reason) one may ask what service would be offered when fares are constrained to be at current levels, but other quality-of-service variables may be adjusted. Fixing minimal levels of these other variables means that the zero-loss position is no longer a lower bound.

Focusing on loss-minimization (or equivalently, on profit-maximization) immediately suggests a way of modeling the structure of competition. For expositional simplicity, consider only a single bus mode and a single BART mode. The profit-maximizing bus level-of-service (including the fare) will depend in the usual way on consumer preferences (demand) and also on the level-of-service offered by the BART system. Suppose that the bus operator takes the BART service level as given. He then sets his own profit-maximizing service level. But of course since BART profits depend on bus service levels, the BARTD will have an incentive to change its offered service. Suppose it does so, taking the previously set bus level-of-service as given. In the next period the bus operator will again change service, reacting to BART service. Eventually (subject to certain technical conditions)[19] an equilibrium may be attained, in which neither agency has an incentive to change further. It is the characteristic of such an equilibrium that will be the object of our study.

This procedure is a simple generalization of a Cournot duopoly. In the usual formulation, each duopolist sets the quantity of a homogeneous product; price is determined by the demand for the total supply. The procedure adopted in the present study is a generalization of the Cournot model in that the products provided by the different transportation agencies are regarded as distinct; and both price and quantity (a vector of quality-of-service indicators) are set by each agency. Thus, the model adopted here may be thought of as a combination of the Cournot duopoly model (in which the choice variable is output) and the Bertrand duopoly model (in which the choice variable is price).[20] In what follows I shall refer to the equilibrium as a Cournot (or Cournot-Nash) equilibrium even though one should, strictly, speak of a Cournot-Nash-Bertrand equilibrium.

1.3 Outline of the Study

The following sections develop formally the model sketched in the last paragraphs. In Section 2 I discuss a model of the demand for the different modes of urban transportation, and show an empirical estimate of the mode-choice model. Since the mode-choice model shows which aspects of service quality are considered important by consumers, its consideration is logically prior to the supply-side discussion. Section 3 develops the Cournot-Nash duopoly framework and sets up the loss-minimizing problem for a combined BART-plus-AC-Transit agency. Section 4 discusses empirical estimates of the supply and demographic parameters. Finally, Section 5 shows simulation results for a wide variety of potential urban transportation scenarios.

2. Demand

In this section I set out the model of mode-choice on which the preferences of consumers for different transit services are based. The model is the familiar disaggregate probabilistic mode-choice model known as the "conditional logit" approach. Section 2.1 contains a brief derivation of the functional form of the logit model; in Section 2.2 I show an empirical estimate of the logit form and discuss its properties.

2.1 Mode-Choice Model

Following the notation of Hausman and Wise (1978), consider a representative consumer indexed by i facing $j = 1,\dots,J$ choices in transportation mode.[21] The consumer is characterized by an observable vector a_i; for example, elements of a_i might include the i-th consumer's income, length of residence in the community, age, or sex. The characteristics of the j-th mode are given by an observable vector X_i^j; here might appear the distance the consumer must walk to access mode j, its costs, travel time, and so on. We suppose that the i-th consumer has a utility function, linear in a vector of unknown weights on the characteristics a_i and X_i^j. Then we may write the utility associated with a choice of the j-th mode as

$$\begin{aligned} (2.1) \qquad U_i^j &= U_i^j(X_i^j, a_i) \quad , \quad \text{all } j \ , \\ &= Z_i^{j\prime}\beta \quad , \quad \text{all } j \ , \end{aligned}$$

where Z_i^j is the numerical concatenation of the elements of X_i^j and a_i ; and β is the unknown weighting vector.

In the real world, of course, we either cannot or do not observe all the factors relevant to the mode-choice decision. We suppose that actual consumers are distinguished from the representative consumer by a stochastic term reflecting these unobserved or unobservable differences in modes or in individuals. If the stochastic term corresponding to the i-th consumer's choice of the j-th mode is ε_i^j , then the utility associated with this choice is

$$(2.2) \qquad U_i^j = U_i^j(X_i^j, a_i) + \varepsilon_i^j = Z_i^{j\prime}\beta + \varepsilon_i^j \quad .$$

The individual is assumed to make the discrete choice which maximizes his utility. Due to the presence of random elements, we can analyze only the probability that a given choice will be made. The probability that the i-th consumer selects the j-th mode is

$$\begin{aligned}(2.3) \qquad P_i^j &= \text{Prob}\left[U_i^j > U_i^1 , U_i^j > U_i^2 , \ldots, U_i^j > U_i^J\right] \\ &= \text{Prob}\left[Z_i^j\beta + \varepsilon_i^j > Z_i^{1\prime}\beta + \varepsilon_i^1 , \ldots, Z_i^{j\prime}\beta + \varepsilon_i^j > Z_i^J + \varepsilon_i^J\right] \\ &= \text{Prob}\left[\varepsilon_i^1 - \varepsilon_i^j < (Z_i^j - Z_i^1)'\beta , \ldots, \varepsilon_i^J - \varepsilon_i^j < (Z_i^j - Z_i^J)'\beta\right] \quad .\end{aligned}$$

That is to say, the probability of selecting mode j is simply the probability that mode j provides more utility than each of the other modes in the choice set. Equation (2.3) can perhaps be made more transparent by observing that is just the joint cumulative distribution function (Ash, 1970) of the J random variables $\varepsilon_i^1 - \varepsilon_i^j$, $\varepsilon_i^2 - \varepsilon_i^2$, ..., $\varepsilon_i^J - \varepsilon_i^j$, evaluated at the point $(z_i^j - z_i^1)'\beta$, $(z_i^j - z_i^2)'\beta, \ldots, (z_i^j - z_i^J)'\beta$.

Any selection of a distribution of the variables ε_i^k will yield functional forms for the mode-choice probability P_i^j. Issues relevant to the selection are discussed at length by Domencich and McFadden (1975): after examining the normal, arctan, and Weibull distributions (all of which yield easily computed selection probabilities in the case of binary choice), they observe that "the multiple-choice generalizations of the [normal] and arctan models are computationally intractable."[22] Thus one is led to characterize the random elements ε_i^k as having independent Weibull distributions, that is,

$$(2.4) \qquad \text{Prob}\,[\varepsilon_i^k < \alpha] = e^{-e^{-\alpha}} .$$

Under this assumption it can be shown[23] that the probability given in equation (2.3) reduces to

$$(2.5) \qquad P_i^j = \frac{e^{z_i^{j\prime}\beta}}{\sum\limits_{\substack{n=1 \\ n\neq j}}^{J} e^{z_i^{n\prime}\beta}} .$$

Observe that this expression involves only the observable elements of choice (and the unknown parameter vector β). A further property of this model, known as the conditional or multinomial logit model, is that

$$\text{(2.6)} \qquad \ln \frac{P_i^j}{P_i^k} = (Z_i^j - Z_i^k)'\beta \quad .$$

That is, the (log) odds of choosing mode j over mode k depend only on the characteristics of the two modes in question. In particular, the odds do not depend on the characteristics of any modes irrelevant to the choice between j and k . For this reason, the property expressed in (2.6) is known as the Independence from Irrelevant Alternatives property (IIA). It is worth emphasizing that it characterizes the logit model; and there may be instances where it cannot plausibly be assumed to hold. A famous counter-example is the so-called Red Bus-Blue Bus Case. Suppose that a population has only two choices available, automobile and a Red Bus. Suppose that twenty percent select the bus. Then the odds (P_a/P_{rb}) of selecting the automobile over the bus are 80/20 = 4 . And now suppose that a second bus is introduced, differing only in some insignificant way from the first--say that it is blue, for example. What we would expect is that the eighty percent of the population now going by car will continue to do so, and that of the twenty percent bus users half will go on the Red Bus and half on the Blue. But the logit model says otherwise. It requires that $(P_a/P_{rb}) = 4$ and that $(P_a/P_{bb}) = 4$ since the two buses are perceived as identical. Then we predict that sixty percent will go by car and that forty percent will go by bus, split equally between the two types. There is clearly something illogical in the result that introducing "more of the same" can induce a large switch to transit. The reason is, of course, that the red bus and blue bus are not independent of each other. The moral of this is that one must be careful to use a logit model only when one has reason to believe that the alternatives are truly distinct. When this assumption is tenable, equations (2.5) and (2.6) show that the mode-choice probabilities are simple functions of the data.

The statistical problem is to estimate the parameter vector β . A random sample of a population characterized by these mode-choice probabilities can be regarded as a drawing from a multinomial distribution with parameters $P_i^1 P_i^2 \cdots P_i^J$. One can then write down the likelihood function of the sample, which will depend on the unknown vector β . Domencich and McFadden (1975) discuss several ways to estimate β : for example, maximum likelihood methods yield consistent asymptotically normal estimates.[24] Since we draw on the work of McFadden and Talvitie for empirical estimates of the mode-choice probabilities, we need dwell no longer on the details.

2.2 Empirical Bay Area Model

A group of researchers at the University of California, Berkeley, has recently completed a major study of the demand for travel in the Bay Area. As part of their research they estimated disaggregate mode-choice probability functions of the form of equation (2.5). I draw on their results to provide the empirical estimates of the models (McFadden and Talvitie, 1977).

Table 3 shows the estimation results of a mode-choice model estimated on a sample of 635 individual decisions in the so-called "post-BART" sample.[25] This contains observations on the seven percent of the population sampled who made part of their trips by BART. The "independence from irrelevant alternatives" property discussed in the last section makes prediction of a new mode's acceptability to the population extremely straightforward, and McFadden and Talvitie do report estimating a mode-choice model from a "pre-BART" sample. For our purpose, which is to discuss the way in which travel demand changes in response to changes in modal organization, it is more appropriate to use the additional information generated by observed choices involving BART. Further, as McFadden and Talvitie note, "the estimates are fairly similar."[26]

TABLE 3

Work-Trip Mode-Choice Model, Estimated Post-BART
with non-generic Auto and Transit on-vehicle time.

(Mode 1--Auto Alone; Mode 2--Bus, Walk Access; Mode 3--Bus, Auto Access; Mode 4--BART, Walk Access; Mode 5--BART, Bus Access; Mode 6--BART, Auto Access; Mode 7--Carpool)

Model: Multinomial Logit, Fitted by the Maximum Likelihood Method

Independent Variable (The variable takes the described value in the alternatives listed in parentheses and zero in nonlisted alternatives)	Estimated Coefficient	t Statistic
Cost divided by post-tax wage, in cents divided by cents per minute (1-7)	-.0266	3.92
Auto on-vehicle time, in minutes (1, 3, 6, 7)	-.0473	3.48
Transit on-vehicle time, in minutes (2-6)	-.0197	2.03
Walk time, in minutes (2-6)	-.0900	3.36
Transfer-wait time, in minutes (2-6)	-.0438	1.81
Number of transfers (2-6)	-.120	0.856
Headway of first transit carrier, in minutes (2-6)	-.0290	2.60
Family income with ceiling of $7,500, in $ per year (1)	-.000289	1.78
Family income minus $7,500 with floor of	.0000522	0.364
Family income minus $10,500 with floor of $0 and ceiling of $5,000, in $ per year (1)	-.0000419	0.738
Number of persons in household who can drive (1)	1.48	5.26
Number of persons in household who can drive (3,6)	1.65	5.16
Number of persons in household who can drive (7)	1.28	4.85
Dummy if person is head of household (1)	.668	3.19
Employment density at work location (1)	-.00164	3.45
Home location in or near CBD (2=in CBD, 1=near CBD, 0 otherwise) (1)	.1546	0.835
Autos per driver with a ceiling of one (1)	4.79	3.70
Autos per driver with a ceiling of one (3, 6)	3.63	4.81
Autos per driver with a ceiling of one (7)	3.26	3.19

Autos alone alternative dummy (1)	-4.18	2.82
Bus-with-auto-access dummy (3)	-8.24	6.67
BART-with-walk-access dummy (4)	-2.28	3.36
BART-with-bus-access dummy (5)	- .473	0.708
BART-with-auto-access dummy (6)	-7.30	5.93
Carpool alternative dummy (7)	-5.31	5.56

Likelihood ratio index	.4599
Log likelihood at zero	-964.4
Log likelihood at convergence	-520.9
Percent correctly predicted	67.24

Values of time saved as a percent of wage (t-statistics in parentheses):

Auto on-vehicle time	178	(2.53)
Transit on-vehicle time	74	(1.84)
Walk time	338	(2.46)
Transfer-wait time	165	(1.65)

Value of initial headways as a percent of wage: 109 (2.13)

All cost and time variables are calculated round-trip. Dependent variable is alternative choice (one for chosen alternative, zero otherwise).

Number of people in sample who chose

Auto alone	378
Bus-with-walk-access	68
Bus-with-auto-access	9
BART-with-walk-access	4
BART-with-auto-access	33
Carpool	137
Total sample size	635

Source: McFadden and Talvitie (1977), pp. 149-150.

The following discussion of the model presented in Table 3 is drawn from McFadden and Talvitie. First, the estimated model allows for a wide choice of transportation modes. If one distinguishes between the linehaul and access components of a trip, and counting as "different modes" trips made on different combinations of linehaul and access modes, then there are in all seven modes in a consumer's choice set. Two involve the automobile alone, allowing for carpooling and driving alone. Two involve the bus as a linehaul mode, with access accomplished by walking or driving to the bus-stop. And three involve BART as a linehaul mode, with access by walk, by bus, or by automobile. The coefficients themselves are generally estimated quite precisely; however, one cannot reject the hypothesis that the number of transfers involved in changing transit carriers has no effect on mode-choice. Three of the demographic descriptors have relatively imprecise estimates: two involve the relationship of family income to mode-choice, and a dummy variable allowing for locational effects is also not well-estimated. Finally, Table 3 reports an imprecise estimate of one of the constant terms.

Turning now to the summary measures of goodness-of-fit, observe first that the model correctly predicted sixty-seven percent of the choices of the sample population. It is instructive to see the gain that this represents over a "chance" model. Without information on the attributes and coefficients of Table 3 each person would be "assigned a probability of taking a given mode equal to the aggregate share of that mode."[27] Under this assignment of probabilities the total percent correct is forty-one percent. Thus, the estimates of Table 3 represent a substantial improvement in predictive power. This is confirmed by the likelihood ratio index. As reported by Domencich and McFadden, this is analogous to the conventional R^2 measure of fit in linear estimation, except that values of the index are downward biased.[28] Thus, a reported likelihood ratio index of .46 represents a better fit than would a similar R^2 index.

McFadden and Talvitie also report on several tests of specification involving the model of Table 3. They reject the hypothesis that, contrary to the model, auto and transit on-vehicle times are valued equally.[29] In the model of Table 3, auto on-vehicle time is valued at 178 percent of the post-tax wage, while transit on-vehicle time is valued at seventy-four percent of the wage. McFadden and Talvitie note that this result runs counter to popular belief about the disutility of transit travel, but observe that the belief may have its roots in consideration of wait and access time as well as on-vehicle time. They test and reject the hypothesis that the relationship between utility and time is nonlinear.[30] However, a modification of the model to allow for differences in travel time and travel cost between urban and suburban dwellers is accepted.[31] This modified model was not adopted in the present research, for two reasons. First, there was a substantial increase in the complexity of the model. An additional five parameters were to be estimated, and to use the model for prediction a further partitioning of the sample by demographic characteristics would have to take place. Second, for the more complicated model, the percentage of actual choices correctly predicted fell slightly.[32] Since the model is used here for explicit prediction purposes, it was decided to retain the model of Table 3.

Finally, it is useful to review the discussion by McFadden and Talvitie of the reasons for mispredictions in the model. They suggest two possible culprits.[33] First, the "independence from irrelevant alternatives" property, a feature of the logit model, may not hold. If the five transit modes were not perceived as independent, then a model requiring them to be so perceived would mispredict the choices of the sample population. McFadden and Talvitie estimate two models without this restriction, and conclude that "it seems that failure of [the "independence from irrelevant alternatives" property] is not a primary cause of the overprediction of transit."

Second, they look to data problems in coding the walk times experienced by consumers using the different modes. These times were calculated from a standard set of network models designed to simulate Bay Area conditions. The reseachers in McFadden's project made various changes to the networks to convert them to 1975 conditions. However, if the original data were faulty, so would all the adjustments have been. In particular, the number of bus lines anticipated to exist may have been higher than the number actually present. This would have the effect of increasing average walk access time. "Secondly, there is often a tendency . . . to code the walk times of those who actually use the system, rather than the average walk time of the population segment. Third and finally, it is possible that some walk times . . . were intentionally coded low in order to predict a large BART patronage."[34]

The upshot of this discussion may be summarized as follows. If the purpose of the mode-choice model is to predict the choices of Bay Area commuters when faced with service variations in the seven modes presently available to them, then the logit model set out in Table 3 would appear to be the best available. It successfully predicted the mode-choices of two-thirds of the sample population, a significant improvement over "chance" predictions. However, the model is not free from defects: in particular, a suspected incorrect coding of walk times may result in overpredictions of the mode-share of Bus/walk, BART/walk, and BART/bus combinations. There would appear to be no way, short of recalculating all walk times in the sample, to remedy the problem. Thus, although the model does represent our best knowledge about Bay Area travel demand, results using this model cannot be expected to be entirely free from error.

The model of Table 3 is used in the sequel to represent the choices of Bay Area commuters. This concludes the discussion of travel demand. The next section begins the specification of the supply side of the model.

3. Supply

This section presents the theoretical model of urban transportation supply sketched briefly in the introduction. The model is completed in Section 4 by empirical estimates of the supply-side parameters. The plan of the present section is as follows. I first discuss the situation of the traveler and argue that it is realistic to model him as having no effect on the conditions facing him should he use his car. I turn then to the issue of competition. It follows from the discussion of the consumer that it suffices to model competition between the providers of transit services. The model used here conceives of two suppliers, a bus company and the BART organization. I show how reaction-function competition may be modeled and discuss the concept of a reaction-function equilibrium. Following this, I discuss the characteristics of an abstract market for transportation, and present a model in which the decision variables of the transit agencies translate precisely into those features of travel revealed as important by the demand models of the last section.

3.1 Consumers

It is clear that we may model the user of the public modes as a price-taker. However, the traveler who uses his car to travel on the existing roadways is in a unique position. He alone has immediate access to his means of transportation; and the consequent ease of travel planning has no doubt contributed significantly to the popularity of this means of locomotion. From the point of view of the economics of his actions, I suggest that they may be characterized by two conditions.

First, it is reasonable to model the automobile user as a "price-taker." That is to say, he perceives the cost of travel as given, and independent of his actions. Whether or not he decides to use his car, the prices he faces will not change. In some important respects this is quite clearly true. For example, an individual's consumption of fuel and oil is likely to be miniscule when compared to the total consumption. Under these circumstances, the effect on the prices of these products of a change in individual travel demand behavior will be negligible, and the assumption of perfect elasticity is wholly reasonable. The same can be said for many of the other variable costs of travel, such as tire costs (per vehicle mile). The assumption is less reasonable when applied to costs like auto insurance; here, particularly, driving patterns do influence the annual premium. Similarly, automobile depreciation costs may vary with the intensity of driving; this might mean, for example, that mileage-related depreciation might be greater for long work trips than for shorter shopping trips. However these variations are likely to be small. Most work on automobile operating costs assumes that depreciation per vehicle mile is constant.[35] Similarly, insurance premiums per vehicle-mile are unlikely to show much variation as between a policy based on auto travel to work and one assuming that work trips are made on transit. Thus the assumption of price-taking behavior on the part of automobile users does not appear to be unreasonable.

Second, I assume that the automobile user faces congestion conditions on the road which are invariant to the traffic on the road. That is to say, a given road is characterized by an attainable speed, and this speed is not related to the volume or composition of the traffic on it. Such an assumption is quite clearly unrealistic, but the following considerations suggest that as an approximation it may be assumed to hold. As will appear in a later section, two types of roads figure in the analysis: expressways (freeways) and arterials. We shall model the trips of concern to the analysis as proceeding from a given origin area to a given destination area. For that portion of the trip taking place on the expressway, it may be assumed that travel between the origin and destination of interest represents a small portion of all the traffic on a given expressway link; for example, traveling on I-80 between Richmond and Oakland one is likely to find a considerable proportion of traffic at a given time originating at, or bound for, other points. Then the effect of a change in the travel patterns of Richmond-Oakland travelers, which represents only a small portion of total I-80 traffic at a given time, will be small in terms of the attainable speed for all vehicles on the link in question.

The question of arterial speeds is less clear and less studied. The explanation given above, that traffic bound for a given destination constitutes a small portion of total traffic volume, is unlikely to hold, simply because of smaller volumes. However, engineering studies suggest that the attainable speed on an arterial is as much dependent on the geometric layout of the road (the frequency of intersections and the provision of traffic lights and stop signs, and so forth) as on traffic volume.[36] To the extent that this is so, an assumption that attainable speed is invariant to volume is more plausible; and that assumption will be made in what follows.

3.2 Competition

I now set out the reaction-function approach to competition between two transit agencies. For simplicity I suppose that to each agency there corresponds a single travel demand. In the mode-choice model presented in the last section this is not so: for example, for the BART mode there are three separate classes of demand to be considered, depending on the access mode. Inclusion of these classes adds nothing but notational complexity; I return to this issue later.

It will be convenient to denote by 1 and 2 the bus and BART modes, respectively. Denote by x^1 and x^2 the vectors of characteristics of the two modes.[37] For example, x^1 might be the bus fare and vehicle headway; x^2 may be thought of as BART fare and headway. Suppose there is a given traveling population of size Q, and that it may be segmented into N homogeneous groups, where the homogeneity is reflected in a socioeconomic parameters vector a_i. Then the mode-choice probability for a member of the n-th group depends on the characteristics x^1 and x^2, and we may write (the precise functional form being given in Section 2):

$$(3.1)\qquad P_n^1 = P_n^1(x^1,x^2,a_n) \quad ,$$

$$P_n^2 = P_n^2(x^1,x^2,a_n) \quad ,$$

(where we have subsumed into the characteristic vectors the weighting vector β of equation (2.5)). If d_n is the proportion of group n to the travel population, then total demand for each mode is given by

$$(3.2) \qquad D^1(x^1,x^2) = Q \sum_n d_n P_n^1 \quad ,$$

$$D^2(x^1,x^2) = Q \sum_n d_n P_n^2 \quad .$$

We can now form and study the profit functions of the two modes. If x_k^1 , the k-th element of x^1 is the fare on model 1 and x_k^2 that on mode 2, and if $C^1(x^1)$ and $C^2(x^2)$ are the total costs of providing service at levels x^1 and x^2 , respectively, then the profit expressions are

$$(3.3) \qquad \Pi^1(x^1,x^2) = x_k^1 D^1(x^1,x^2) - C(x^1) \quad .$$

$$\Pi^2(x^1,x^2) = x_k^2 D^2(x^1,x^2) - C(x^2) \quad .$$

We model each transit agency in behaving in a Cournot-like way. Each takes the service quality vector of the other as given, and maximizes profits. This leads to a set of reaction functions, one for each agency, where the reaction function indicates the best choice of one's own level-of-service vector conditional on that of the other. If $\bar{x}^1$ and $\bar{x}^2$ are arbitrary fixed levels of service for modes 1 and 2 then the reaction functions are defined by

$$(3.3a) \qquad R^1(\bar{x}^2) = \max_{x^1} \Pi^1(x^1,\bar{x}^2) \quad ,$$

$$R^2(\bar{x}^1) = \max_{x^2} \Pi^2(\bar{x}^1,x^2) \quad .$$

A reaction-function equilibrium is a pair of vectors $\hat{x}^1$ and $\hat{x}^2$ which satisfy

$$(3.4) \qquad \hat{x}_1 = R^1(\hat{x}_2) \quad ,$$

$$\hat{x}_2 = R^2(\hat{x}_1) \quad .$$

This reaction function equilibrium will be the object of our empirical study in Section 5; it represents a stable point in the sequential reactions of one agency to the actions of the other. It is therefore worth pausing to consider two issues related to the formulation.

First, as has been commonly observed, there is something implausible in the assumption that the behavior of each competitor results from considering as given the actions of the other. At each stage of the competitive process each expects the service quality of the other to remain fixed; but each time (until the attainment of equilibrium) the expectation is not realized. One would surely expect some learning to take place, and that each competitor would realize the folly of assuming that the other's service was fixed.

There are indeed incentives for this to happen. Suppose that the first mode knows the reaction function of the second. Then rather than maximizing

$$\Pi^1 = x^1_K D^1(x^1, \overline{x}^2) - C(x^1) \quad ,$$

and moving sequentially towards a reaction-function equilibrium, it is much more reasonable to maximize

$$\Pi^{*} = x^{1}_{K}D^{1}(x^{1},R^{2}(x^{1})) - C(x^{1})$$

--that is, to take explicit account of the reaction of the other. It can be shown that this sort of competition-- known in the single-output case as Stackelberg competition-- generally results in higher profits for the leading firm.[38]

Three reasons may be adduced for not modeling competition in this way. First, there is a conceptual difficulty if both firms try to behave in this way. An equilibrium may not exist at all, and there may be constant movement around disequilibrium positions. Second, as the number of steps to the reaction function equilibrium decrease, there is less room for learning, and a smaller chance of the anomaly of continually unfulfilled expectations being noticed. As will appear in Section 5, the number of steps to a reaction-function equilibrium is generally small; hence one may expect learning behavior to be less important. The third difficulty is one of implementation. The Stackelberg solution requires that the reaction functions be explicitly incorporated into the demand functions. It will be shown in the next section that an explicit solution for the reaction functions is not possible; rather, recourse must be had to numerical algorithms to solve at each stage the pair of problems (3.3). Thus it is empirically infeasible to model this concept of competition.

The second issue related to the reaction function formulation of the problem is that it might be supposed that one set of actors has been left out entirely. The discussion thus far has focused only on the transit carriers; no mention has been made of the competition from the automobile mode. However, this is wrong: consideration of the transit modes alone suffices to characterize the equilibrium. The reason lies in the demand functions for transportation modes, or, more specifically, in the mode-choice probabilities which form the basis for computing the total revenue of the transit modes. As we saw in Section 2.1, these are explicitly derived from considerations of utility maximization. Thus at each stage of the search for reaction-function equilibria there is built in to the model an implicit utility maximization on the part of price-taking consumers. Thus there is no need to take explicit account of consumers at all.

3.3 The Transportation Market

In this section I place the reaction-function competition described in the last section in the more detailed context of an abstract transportation market. This is shown schematically in Figure 1.[39] The structure of the market is one of transportation between a set of origins and a given destination area, such as a CBD or major shopping center. In the Bay Area, the type of market envisaged in Figure 1 might be commutation between Richmond and Oakland or San Francisco,[40] or between Orinda-Lafayette and the same CBDs.

The abstract structure of the market is as follows. One has a circular residential area of radius r miles. We suppose that origins are uniformly distributed over the radii of the residential circle; thus the "average" resident lives $r/2$ miles from the center. We assume too that the relevant socioeconomic characteristics (to be described in Section 3.4) are also uniformly distributed. Thus we rule out the possibility of (say) the left side of the residential area being poor and the right side being rich. In this formulation rich and poor live together.[41]

Trip-making between origin and destination takes place as follows. Suppose the trip-maker lives at H . If the selected mode is the automobile the consumer drives from H to A , which has the interpretation of a freeway entrance. He drives at freeway speeds from A to C , the freeway exit, and then proceeds along arterial streets from C to D to E , where he parks. From E he walks to work at W, assumed to be .25 miles away from the parking lot. If he carpools, it is assumed that this results in a "carpool delay" at the residential location: the rest of the trip is as given previously.

The description of bus trips is slightly more complicated. From home at H the commuter either walks or drives to the nearest bus stop, assumed to be located at G. When the next bus arrives the commuter boards it; the bus proceeds along arterial streets to B, which is also a freeway entrance. Note that this formulation introduces a perceived disadvantage for the bus, equivalent to the fact that the bus takes a more circuitous route to the freeway. The present structure, instead of increasing bus distance traveled introduces a circuity factor by having the bus travel from A to B (a distance of XL/2 miles) at the slower arterial speed. The automobile user traverses the same distance at the faster expressway speed. Note that this adequately captures the circuity-induced disadvantage since modal choices depend only on travel time and not distance (from Table 3). The bus proceeds on a freeway from B to C, proceeds on arterial roads from C to D to E; at E the commuter alights and walks to work at W.

Trips by BART are again slightly different. Here we have three possible access modes. Assume that the BART station is located at A. If the access mode is the automobile or by walking, the consumer proceeds directly from H to A. If the access mode is automobile, it is assumed that free parking is available at or immediately adjacent to the BART station. If BART is accessed by bus, then again he walks to the nearest bus stop, waits for the bus, and rides on a bus to A . At the station he waits for the next train, which takes him from A to E . Note that if the BART speed is high enough, this represents an advantage in travel time over the competing modes which travel at arterial speeds for XL/2 miles (auto) or XL miles (bus). The BART station at the destination is assumed to be located at E : from there one has the usual walk to work of W .

I turn now to a discussion of the concept of transit service quality in this model. The demand model of Table 3 shows that consumer preferences are based on the following measures: cost, on-vehicle time, walk time, transfer wait time between transit modes, number of transfers and headway of first transit carrier. From the assumptions discussed previously, highway speeds, and hence travel times, are given. I shall make a similar assumption that BART speeds are likewise given. The structure of the market determines the number of transfers between transit modes. Hence, the measures which are subject to agency determination are fare, walk time, transfer wait time, and headway of the first transit carrier. We shall incorporate the last three in the form of operating characteristics.

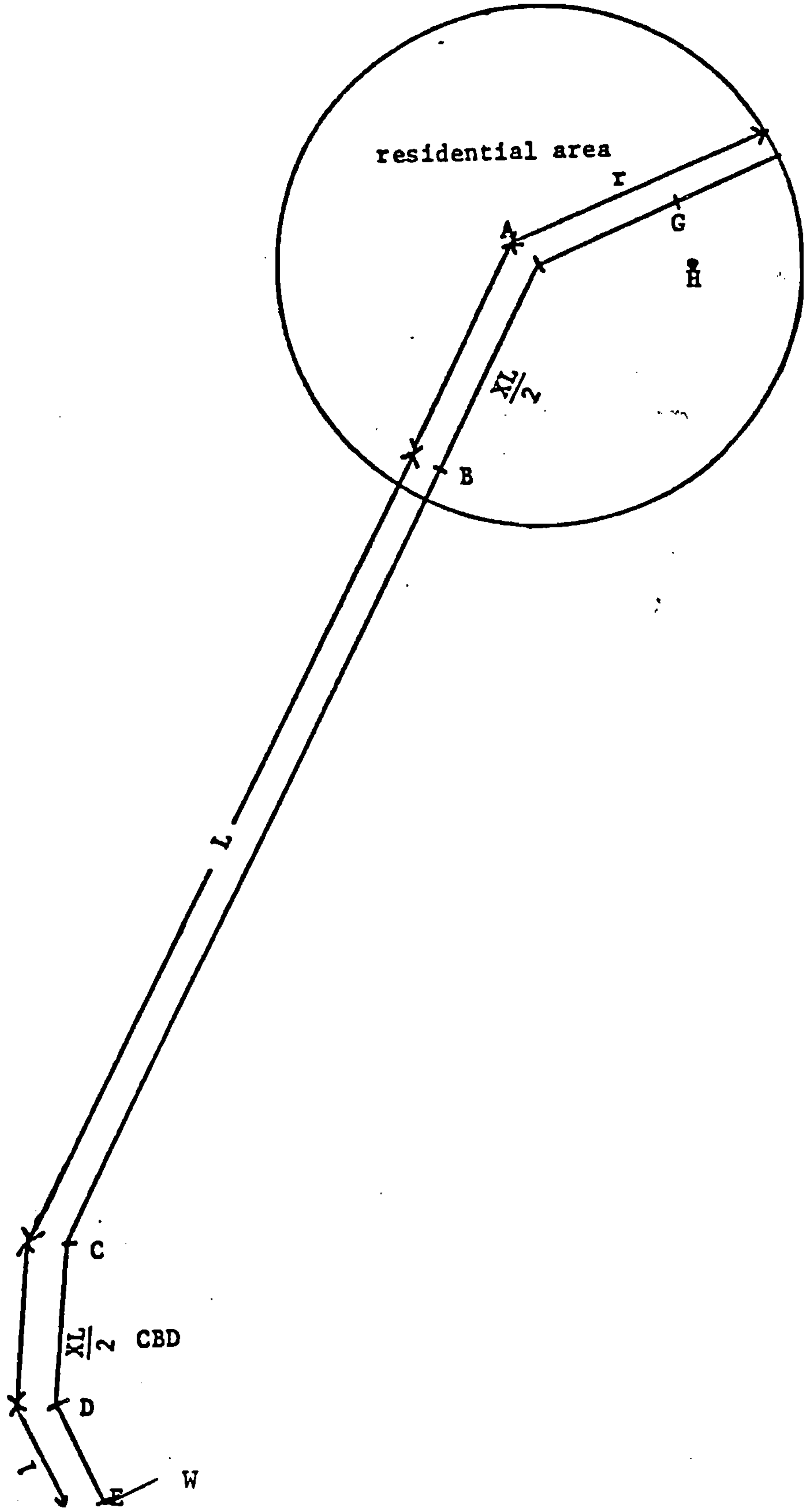
We now make the first major organizational assumption of the model. Since we are interested in describing the outcome of a BART/bus access system run to feed efficiently to the BART system, we suppose that there is a feeder bus service distinct from the regular bus service, and that it is run by BART only to feed into the rapid rail system. Thus we conceive of a BART system which also runs its own exclusive feeder buses for those who choose to get to the trains by bus. This means that along some of the residential streets there will be in effect two bus lines: one destined for the CBD and one for the nearest BART station. We allow the bus systems to be entirely separate, running at different levels of service and charging different fares.[42]

Suppose that the bus system (BART feeder system) runs R (R_f) equally spaced bus routes over the residential area. From previous assumptions, the average commuter lives $r/2$ miles from the freeway entrance/BART station at A in Figure 2. At this level of bus services, the average commuter, who is assumed to walk at three miles per hour, will take $\pi r/6R$ ($\pi r/6R_f$) hours to walk to the nearest bus (feeder bus) stop. Thus, from the description of the level at which the residential area is "covered" with bus routes we can immediately deduce the average access walk time a commuter faces. (Recall that the location of the BART tracks is fixed.)

Similarly, denote by ψ (ψ_f, ψ_B) the headway of buses (feeder buses, BART trains) along their routes. If arrivals at bus stops, (feeder stops, BART stations) follow a Poisson distribution, then the relation of average wait at a stop is simply "wait equals half the headway." Fisher and Viton (1974) have investigated other distributions of arrivals under which this result holds approximately: they conclude that for short headways (less that ten minutes), this rule provides an accurate representation of the true wait.[43] So from the system variables of vehicle headways, we may deduce the average wait times. Then the results of the last paragraphs imply that by considering only cost (auto cost, transit fares), number of

Figure 1

Transportation Market



routes in the residential area provided by the "integrated" and "feeder" bus systems, and the headways of the transit modes we can capture all the relevant measures of service quality revealed as important by the choice model of Table 3.

Finally, we restrict the mode-choice model to a six-mode model by assuming that the extent of carpooling, represented by an average number of travelers per car, is fixed and not an object of choice by the auto-using segment of the market.

3.4 Mathematical Formulation

We now set up the structure from which we may write the detailed equations for the profit functions of the bus and BART operators. We index the modes by their indices in Table 3: thus, for example, P_n^4 is the probability that a member of the n-th homogeneous group of travelers will choose to travel by BART, after walking to the system. The last section showed how the independent variables in these probabilities are related to the system characteristics. Again, we assume that the total traveling population in the period under notice is Q .

First, we set up the profit expression for the bus system. If, as we suppose, each rider is charged an identical fare f , then the bus operator chooses R , ψ , and f to maximize

$$\Pi = Qf \sum_n (P_n^2 + P_n^3) d_n - C(\psi,R) \quad , \tag{3.5}$$

where $C(\psi,R)$ is the cost of providing this level of service.

The BART operator provides in effect two modes: the BART rapid-rail system and the feeder bus system. Denote by f_B and f_f the BART and feeder fares, by R_f the number of feeder routes, and by ψ_B and ψ_f the headways of BART trains and feeder buses. Then the BART agency selects f_B , f_f , R_f , ψ_B , and ψ_f to maximize

$$\Pi^B = Qf_B \sum_n \left[P_n^4 + P_n^5 + P_n^6 d_n\right] + Qf_f \sum_n P_n^5 d_n - C^B(R_f,\psi_B,\psi_f) \tag{3.6}$$

where C^B is the cost of providing the combined level of service. This formulation is deceptively simple: recall that the probabilities P_i^j are given by equation (2.5), and they are each non-linear functions of all the choice variables. Writing out the necessary conditions for maxima of equations (3.5) and (3.6) is a complicated and tedious operation: no point would be served by the exercise since it is clear from an inspection of the equations that a closed-form solution is not possible.

Finally we impose a set of constraints on each of the profit-maximization problems. There is an upper limit to the number of persons who can enter a single bus; we take this to be the seating capacity (fifty) of a bus. Then we require that

$$(3.7) \qquad \psi Q \sum_n (P_n^2 + P_n^3) d_n \leq 50 \quad ,$$

$$\psi_f Q \sum_n (P_n^5) d_n \leq 50 \quad ,$$

for the regular and feeder systems, respectively.

For BART there is also a capacity constraint, though it depends on the number of cars in a train. If T is the number of cars, and we require that each car hold not more than 100 commuters:[44]

$$(3.8) \qquad \psi_B Q \sum_n (P_n^4 + P_n^5 + P_n^6) d_n \leq 100T \quad .$$

There is also, in the case of BART, a safe-headway requirement. Because of the confluence of three lines at Oakland, trains there must be 120 seconds apart, or six minutes on each of the three lines. Therefore on the BART line we require

$$(3.9) \qquad \psi_B \geqq .1 \quad .$$

A complete listing of all the detailed equations is in Appendix A.

The cost terms are discussed in the next section. Though the model formulation is quite general, the following additional assumptions will be made. The demand model of the last section is for work-trips; hence attention is focused on the peak period alone.

It is assumed that all peak travel is time-of-day inelastic, that is, that there is no possibility of a peak traveler choosing to travel instead in the off-peak, when conditions of service may be more congenial. Second, the influence of minor-direction traffic is ignored. In light of the focus on the peak alone, this is a reasonable assumption, especially for CBD-bound travel. For other periods it is of course questionable; however, given the fact that off-peak travel volumes are low compared to peak volumes, the results in Section 5 indicate that the possibilities of covering variable costs in low-demand periods are quite limited.

3.5 Other Considerations

The formulation thus far has modeled reaction-function competition between the two modes. However, as noted in the Introduction, several other questions are of interest. First, as has been observed, with system parameters freely variable, the worst any transit system can do is make zero profits. By shutting down operations altogether, a loss can be avoided. But one may also think of public transit as being required to maintain a given level of service, as described by the fare, vehicle headway and (except for the BART trains) degree of system accessibility. With these constraints one can no longer be sure the system will no longer incur a loss. In terms of the formulation given above, we need merely incorporate additional constraints of the form $f \leq f^*$ — i.e., that bus fare be less than some determined maximum f^*. The same form is adopted for other system constraints.

Second, the issue of a transit "super-agency" may be approached by examining the profit expression of a transit monopolist. The monopolist will select all service characteristics to maximize total profit $\Pi + \Pi^B$, as given by the sum of equations (3.5) and (3.6), subject to (3.7) - (3.9). In this problem, no reaction-function considerations appear, since the auto mode is a (competitive) residual category.

4. Empirical Parameters

This section sets out the values of the cost and demographic parameters used in the sequel. A number of treatments of Bay Area transportation costs have appeared, and these are relied on heavily. Likewise, for the demographic parameters, data from various censuses are used.

4.1 Transportation Costs

Table 4 compiles the costs of transportation in the Bay Area for the three modes of interest. For the automobile I use as an exemplar a compact car. It would be possible, though notationally difficult, to consider different sizes of cars and then assign them to the different demographic groups; but this is not done here. The auto costs include only the direct out-of-pocket costs, in order to make them compatible with the way in which McFadden's researchers calculated them.[45] No external costs are assumed: this is in keeping with current practice, though taxes, which are included, do cover most of the non-congestion external costs.[46]

TABLE 4

Transportation Costs[1]

1. Automobile--Compact Car	
Arterial Cost/vehicle mile[2]	$.0827
Linehaul Cost/vehicle mile[2]	.0597
Parking Cost/day[3]	1.71
2. Bus	
Cost/major direction peak vehicle-hour[4]	16.34
Cost/major direction peak vehicle-mile[5]	.3274
3. BART	
Cost per peak car-mile[6]	7.27
Current cost/car-mile	2.85
Vehicle Capital cost/car-mile	4.42

Notes

[1] The references below give costs in 1972 dollars. The figures in the table multiply those costs by 1.3516, the 1972-1976 increase in the San Francisco-Oakland consumer price index (Statistical Abstract of the United States, 1978, Table 772).

[2] Keeler *et al.*, 1975, pp. 84-86. Includes only costs of auto operation and maintenance.

[3] Rough average of all-day rates in 1978 for seven downtown Oakland lots.

[4] Viton, 1977, p. 50.

[5] *Ibid*. Includes $.0737 administrative expenses.

[6] Keeler *et al.*, 1975, p. 122. Six percent interest rate assumed. The next two lines show the cost allocation between current and vehicle capital costs.

It is assumed that both the regular bus system and BART's feeder bus system use the same technology, and are subject to the same costs. In at least one instance we observe feeder buses which are substantially smaller than regular transit buses, but this innovation is not discussed.[47] All capital costs of buses are assigned to a four-hour daily peak, and are included in these costs.

BART costs include the average costs of running the system, plus a capital cost of the rolling stock. Thus it is assumed that in the short-run rolling stock costs are variable: that is, that the fleet could be easly disposed of. By the same reasoning, all other capital costs are excluded: it is assumed that the capital equipment (excluding rolling stock) is fixed, and that its costs would continue to be incurred if no output were produced.

The model here implicitly assumes constant returns to density over the relevant portion of the BART cost curve and the bus cost function. The constant-costs assumption for buses is standard in the literature. Pozdena and Merewitz present evidence that this is not so for rapid-rail: in particular, they estimate that rapid-transit properties show economies of density in the short-run. However, even in their Model 4, one cannot reject at the fifteen percent level of statistical significance the hypothesis of constant returns to density.[48] Thus, in what follows, constant returns in the short-run are assumed; and average and marginal costs are the same.

All costs are converted from 1972 to 1975 dollars. Clearly costs have changed since 1975, but no attempt is made to take account of them. Thre are two reasons for this. First, while the simple adjustment procedure adopted here will be reasonable for the period 1972-1975, it is unlikely to be accurate for a longer period. To redo the results of the Transportation Cost Study (Keeler, _et al_., (1975)) is not practicable. Second, insofar as gasoline costs are perceived as the greatest change since 1972, Keeler, _et al_., suggest that the per-trip costs will not be particularly sensitive to this item.

4.2 Physical Parameters

Table 5 summarizes the physical parameters to be used in the sequel. Note that the average trip length in this formulation is seventeen miles; this is slightly longer than the average Bay Area trip length of sixteen miles.[49] The trip described here may be thought of as an approximation to Richmond-San Francisco travel; though in the abstract-market framework adopted here, this interpretation is not necessary. Recall that we assume that speeds on any road segment are invariant to the volume of travel on that road;[50] see Viton (1977) for a simple model in which speed does depend on traffic.

TABLE 5

Transportation Market Parameters

Symbol	Segment	Reference in Fig. 1	Distance (mi)	SPEEDS Auto	Bus	BART
L	Linehaul	BC	12	40	40	--
r	Residential	AI	2	25	12	--
XL	Circuity	AB, CD	4	25	12	--
	BART Line	AE	17	--	--	45

Transfer Times	Minutes
Bus to BART (150 yds, 3 mph)	1.7
Auto to BART (100 yds, 3 mph)	1.1

4.3 Demographic Characteristics

Table 6 lists the demographic variables used in estimating the mode-choice probabilities, and the distribution of Bay Area incomes. Pre-tax income, household size, and autos per household, and their distributions, are based on the 1975 Survey of Housing and Transportation conducted jointly by the U.S. Department of Transportation and U.S. Department of Housing and Urban Development. The post-tax wage was constructed using average Federal and California tax rates,[51] and assuming a forty-hour work week and a fifty-week work year.

A rough estimate of the number of drivers per household may be obtained by assuming that the household contains two adults and then estimating the probability that the remaining members are over sixteen (and hence eligible for drivers' licenses). Based on 1970 data, forty-six percent of children were over sixteen in 1970[52] and based on data for the California population we take the probability of having a driver's license, given that one is of age, to be .9 .[53]

Because of the aggregation involved in constructing the data, it must be considered as an approximation to the actual situation: however, the socioeconomic data are of relatively less importance in determining mode-choice probabilities than the characteristics of the transportation market.[54]

5. Competition

I now present some results of the inter-modal competition model for several cases of interest.[55] The approach taken here is to make the size of the market (the number of trips from origin to destination in a given peak hour) a parameter; and then to discuss the competitive possibilities dependent on that market size. Recall that the model is a route model: with three East Bay routes, a given market size is equivalent, assuming equal demands on each line, to a system-wide market three times that size. Also note that, although the term "competition" is used, in fact there can be a substantial exercise of market power. Each of the two transit modes perceives a downward-sloping demand curve, and, with price-taking consumers, sets fare and level-of-service to maximize profits. To the extent that this results in large profits, some sort of taxing or regulatory authority would be needed to balance social and private benefits. However, this question is not discussed here. I focus only on the nature of the competition between the two operators, and leave open the knotty question of the distribution of gains to market power. Note too that we focus on peak travel: to the extent that off-peak service is required, peak-period profits might be used to subsidize it.

It is worth re-emphasizing that this is a simulation model, designed to answer a series of "what-if" questions. Those answers depend on the driving assumptions of the conceptual framework. The crucial behavioral assumptions are that at least one of the transit firms try to minimize losses; and that they react to each other's decisions. Thus, when, as will be seen, we conclude that there are significant possibilities for BART's covering its variable costs, it must be recognized that this depends on the BART management acting as reacting loss-minimizers. Clearly, there may be reasons of public policy why this assumption might not be desirable in actual operations--for example, there may be a political or social bias against the concomitantly necessary high fares. But it is important to distinguish this type of reason for subsidized service from the implicitly-made claim that rapid-rail services cannot cover costs. Both the results of the simulation model and at least one actual example--the early history of the Lindenwold Line in Philadelphia (Morlok and Viton (1980))--suggest that this latter claim is wrong. It would therefore appear that the focus of policy discussions should shift from questions of whether BART <u>can</u> cover its costs to whether it <u>should</u> do so.

Finally, two more technical comments. Recall that we shall discuss only equilibria, and not the interim points on the reaction functions. Roberts and Sonnenschein (1977) have pointed to the possibility of multiple equilibria in complex models; and indeed, no proof of uniqueness is attempted here. However, in light of the findings below, it is the <u>existence</u> of profitable equilibria that is important. Second, in all cases reported here, the rate of convergence was sufficiently fast--at most four iterations were required--to make plausible the Cournot assumptions (see Section 3.2).

5.1 Fixed Ten-Car BART Trains

BART's original plan for full operation called for the scheduling of ten-car trains in the peak, and five-car trains in the off-peak.[56] By 1975 this level of operation had not been reached: as Merewitz (1975) shows, the average peak train size ranged from seven cars on the Concord-Daly City line to four in Richmond-Fremont service. This level of service appears to be partly the result of the delivery of defective cars, and the BART District initiated a legal action for damages against the car manufacturers, which has now been resolved in BART's favor. The eventual aim is still to run ten-car peak trains; and therefore this case is worth analyzing.

Table 7 summarizes the outcome.[57] If the aim of BART were to minimize loss with no constraints on the level of service (and hence zero variable costs are incurred at zero output) then only at high levels of corridor travel--approximately 10,000 per peak hour or more--should service be provided at all. In cases where travel volumes range from 2,500 to 7,500 per peak hour, only the integrated (express) bus system covers its costs. In these situations, the best strategy is to close down BART operations. It is true that the capital costs of the system will still be incurred; but at least no operations losses will accrue. Finally, when travel volumes are low--1,000 per peak hour or less--no public mode is able to cover its variable costs. This is in line with a previous result of Keeler, _et al_.--that at low volumes, the full costs of operating a private car are lower than those on an integrated bus mode.[58] However, that study did not explicitly consider the nature of the competition.

In the single case where BART is competitive with the bus service and the private car, it attains a respectable thirty-two percent share of the market, while the bus system carries thirty-seven percent. As to the level of service offered by the two public modes, we obseve that the profit-maximizing BART fare is actually less than the fare set by the express-bus operator. The reason has to do with the constraint imposed on the BART system, that it run ten-car trains. With this constraint, the average walk plus wait time of a commuter on the BART system with 10,000 peak-hour travellers is fifteen minutes; by contrast, the more flexible level of capacity provided by the integrated express bus system gives bus users an average wait plus walk time of fourteen minutes. The lower BART fare compensates for the higher user (time) costs.

Observe, too, the nature of the BART feeder service. With equipment type and costs identical to those of the integrated bus system, the feeder is designed to provide substantially better service. The feeder system operates twenty-six routes in the residential area, while access to the integrated bus system is over eleven routes. Feeder headways are seven minutes, as compared to sixteen on the regular bus system. Finally, the feeder service is provided free: in 1975 a commuter using bus access to BART paid an average fare of $.125 .[59] The result of this efficient feeder service is apparent in the choices of the consumers. The BART plus feeder system carries thirty-two percent of the commuting population: of these, twenty-four percent get to the trains by feeder bus, and seven percent use a kiss-and-ride access mode.

In this case, the extent of product differentiation is not large. Both public modes provide expensive but high-quality service, at roughly comparable travel times and fares. In this respect the equilibrium is not unlike the qualitative situation now observed (recall Table 1). But of course to cover costs, both modes exploit their market power in the equilibrium presented here.

5.2 BART with Free Choice of Train Size

The results of the last section indicate that BART faces very limited possibilities of covering its variable costs when it is constrained to provide service by ten-car trains. The fact that the stations were built to handle long trains, however, should not blind us to the possibility of running shorter tains, as indeed is now being done. In this section we ask if there might not be considerable virtue in the necessity, and whether the eventual operating policy as originally foreseen might not profitably be changed. For these purposes, we assume that BART can freely select the size of its trains, constrained in only three respects. The first constraint, a technological necessity, is that trains have at least two cars. The second constraint is that each car have in it no more than 100 persons. As previously noted, this load factor is derived on a space-equivalent basis: at 100 persons per BART car, each commuter has space equivalent to that he would occupy on a bus seat. The third constraint is also technological. It is assumed that, no matter what size trains are run, there must be a minimum of 120 seconds between trains as they go through Oakland. This is equivalent, with three lines feeding into Oakland, to minimum headways of six minutes on each line. It is possible that safe headways would fall as train size decreased, but this is not investigated here.

The competitive possibilities are shown in Table 8, and they are striking in that, while with ten-car trains BART could only cover variable costs at travel densities of 10,000 per hour or more, when the system can adjust the size of its trains, it can compete at any travel volume exceeding 2,500 commuters per hour. Just on this basis, then, one may conclude that BART would be ill-advised to insist on ten-car train service.[60] It is clear why there is a gain to such a policy. With ten-car trains, BART would be forced to wait for more commuters to arrive at a station, in order to reduce excess capacity. But this imposes a greater time burden on commuters, and they are less willing to use the system. Comparing Tables 7 and 8, and assuming 10,000 commuters per peak hour, we see that when operating ten-car trains BART runs at a headway of nineteen minutes, while with its best choice of train size headways are six minutes. Note, too, that people appear to be willing to pay for the increased convenience: the fare charged goes from $2.03 (ten-car train) to $3.39 (variable train size).

The BART plus feeder system always chooses to provide a free feeder service to the trains; and this is provided at headways less than those provided by the competitive bus system. The BART system, too, provides significantly greater coverage of the residential area than does the bus system--in the case of 10,000 hourly commuters, twenty-five routes are provided by the feeder system, as opposed to sixteen by buses.

In these cases product differentiation takes place in that the combined BART plus feeder system is perceived to provide better service at a higher price. Comparing the two at a travel volume of 10,000 commuters per peak hour, access time to the bus system is twenty-five minutes; while to the feeder system it is 4.9 minutes. Average waiting times are 5.5 minutes for the express bus option, and seven or three minutes for BART, depending on whether or not the feeder system is used. What this means can be seen in the distribution demands over the various income groups.[61] If we look at the lowest and highest income classes (groups 1 and 5) the following emerges. First, auto use differs considerably: thirty-nine percent of group 1 use the car, while sixty-two percent of group 5 do so. Substantial differences emerge in transit use as well. The low-income commuters overwhelmingly use the bus--seventy-eight percent of transit travel takes place on this mode, with virtually all (ninety-six percent) walking to the system. By contrast, sixty-five percent of transit travel by the high-income commuters is made on BART, with only thirty percent of that population electing to use the feeder service. With high values of time, the members of group 5 tend heavily to use their cars to get to the BART system.

These differences in travel patterns are also reflected in the fares charged. The bus fare is $2.30, while the BART system charges $3.39 . Clearly, the wealthy consumers are willing to pay for the superior service option BART offers. This is in line with the empirical examples discussed by Morlok and Viton (1980).

And one conclusion stands out clearly. If it is desired to operate BART in such a way as to cover variable costs (and even make a contribution, through the exploitation of market power, to capital costs) then it appears best to regard the station size as sunk investment, and run small trains frequently. And if this is done, cost-covering operations are feasible even at quite low corridor travel volumes.

5.3 Quality-of-Service Competition

It will be observed from Tables 7 and 8 that where BART is able to cover its costs, it does so by charging an extremely high fare. When unconstrained by the necessity of running ten-car trains, BART fares vary between $2.23 and $3.39 per trip. This is substantially more than originally planned. For example, for the trip described here, the initial (1972) fare structure for BART, as discussed by Pozdena (1972) would result in a fare of $.75 . At the fare levels prevalent in 1975/76 an express bus trip would cost $.65 . Evidently, both modes are exploiting their market power in the model outcomes.

Inter-modal competition can be thought of as occurring in two parts: fare competition and service quality competition, where service quality is defined to exclude the fare. In scheduled trunk airline service, for example, regulation by the Civil Aeronautics Board has severely limited fare competition, and the result has been that the competitive struggle has been expressed in service-quality competition (Keeler, 1977). There are even results suggesting that with three or more competitors in a given market, service-quality competition with fares substantiallly above marginal cost may lead to too good a service prevailing (Eads, 1975).

There is thus some precedent in the transportation industry to look at service-quality competition with fixed fares. In the case of the two competing public modes of interest in this study, the question is relatively easy to resolve. At fares at their 1975/76 levels (integrated bus, $.75 ; BART, $.65 ; feeder bus, $.13) no mode can cover its variable costs; and hence, under the assumptions underlying the model, would all shut down, leaving the private car the only available mode. In practice, this conclusion may be slightly modified by considering minor-direction traffic: recall from the discussion of Section 3 that all costs are assigned to the major direction, and that there is essentially no travel in the minor direction. Under these assumptions, one may calculate the minimal fares necessary to cover variable costs, at a full load: bus, \$.83 ; BART, \$2.47 ; and feeder bus, \$.14 . But to the extent that service competition is to take place, fares must be maintained at levels much above these, since service competition would mean running at less than full loads.

5.4 Unprofitable Bus System

As against the results thus far presented, it can be argued that they do not fairly present the potential for BART, for two reasons. The first is that the model so far discussed presents the integrated bus systems as a profit-maximizing duopolist, and this does not adequately capture the motivations of AC Transit. The second reason is derived from the first: when BART was being planned, it was reasonable to view the possibilities of covering variable costs in the light of the prevailing system of bus transit in which good but money-losing service was the rule. As pointed out in Section 1, however, BART has not in general moved to exploit any opportunities here, since service levels are quite similar to those of the competing linehaul bus. The question of what BART could do in the face of a linehaul bus system with no cost-covering constraint is nonetheless interesting.

With free choice of fares and service levels, the bus system's best action is to shut down when variable costs cannot be covered. This does not hold when the choice variables are constrained. In that case, satisfaction of the constraints may be possible only by losing money. In that case the model gives the loss-minimizing level of service, subject to satisfaction of the constraint.

To discuss a constrained bus system we conceive of two sets of constraints. The first is the fare. To the extent that the BART planners took the competing bus system into account, they probably did so in such a fashion as to assume that the basic low-fare service would continue. The second set of constraints directly affects the quality of bus service offered. Rather than put constraints directly on the two remaining choice variables (bus headway and route coverage) we impose a user-perceived service constraint: that average wait plus access time be less than some specified amount. This leaves the bus system free to adjust headways and route coverage to lower costs as much as possible, just as long as this basic service is provided.

We therefore make two assumptions about the nature of the bus system. The first is that the fare be set at 1975/76 levels, that is, $.65 per rider. We also assume that the average wait plus access time for bus users must be less than twenty minutes. The question of interest is whether BART can still cover variable costs, given this bus competition; and we assume that BART can freely select its train size.

Table 9 shows the result. Two general features are of interest: first, the bus system always provides service, though it never covers its costs. Second, as far as mode shares are concerned, the bus system carries a significantly larger share of the traffic. And although the bus service characteristics are very different from those of Table 8, the BART service characteristics are much the same.

Turning now to the details of the service, recall that the bus system is in fact operating under three constraints, the third being the service constraint common to all the problems studied, that all passengers be seated. This, in conjunction with the other two ($.65 fare, wait plus access time less than twenty minutes) accounts for the rather different service characteristics at different travel volumes. As the number of travelers increases, it becomes progressively easier to run full buses and maintain the required level of service. At large travel volumes the number of people who want to ride the $.65 express bus is large: to accommodate them seated, many buses must be run. But one can reduce the number of routes to make service less attractive, and this is what happens.[62]

The BART and feeder service is unremarkable. At all travel volumes over 2,500 per hour BART is able to distinguish its high-fare service from the low-fare, relatively time-costly bus service, and cover its costs. Again the feeder service is free; BART's share is between twelve and twenty-six percent.

5.5 A Single Transit Agency

Up to now we have modeled the public modes as two competing systems. It has often been suggested to merge the agencies into a single Bay Area transit agency, and the loss of deficit financing attributed to Proposition 13 in the case of the bus system has given added urgency to this question. It is of interest to examine the price and level-of-service which would be offered by such a super-agency and to compare them with those offered by the separate agencies competing against each other. We assume in this section that there are no administrative economies of merger.

Qualitatively it is easy to say what ought to occur. Since the single agency would possess more market power than the separate transit operators, one would expect it to use this power to raise fares and cut service quality, just as a classical monopolist offers less output at a higher price than an industry made up of competitors. What we want to know in the present context is how dramatic a reduction in service would be instituted by the super-agency.

To answer this question we examine the product-choice decisions of a transit monopolist. The monopolist selects three headways (bus, feeder, and BART); designs two route systems (bus and BART); chooses three fares, and the size of the BART trains to maximize total system profits. These are given by the sum of (3.5) and (3.6), and the monopolist is as usual constrained by (3.7)-(3.9).

Table 10 gives some representative answers for the situation in which all fare/service quality variables may be freely set. The following conclusions follow from that table: these conclusions, insofar as they are qualitative, hold for all cases of competition examined. First, as expected, service quality on both modes suffers: in the case of 10,000 hourly travelers in the market, the BART plus feeder service results in average wait plus walk time of ten minutes, as compared with nine minutes in the competing case. For bus service, under the super-agency, average wait plus walk time is eleven minutes, as compared to six under competition. Second, and perhaps more important, the fares on all modes rise. The integrated profit-maximizing bus systems's fare increases by \$.59 , while the BART plus feeder bus system fare rises \$.87 . Not surprisingly, auto travel increases: the transit mode share falls from sixty-five percent when the two modes compete to fifty-two percent under a super-agency.

The same results hold qualitatively for the other travel volumes of Table 10, as well as for the other cases of competition examined previously. In no case would a super-agency close down one system when competition between the two separately is feasible. By the same token, in none of the cases would a mode which under competition could not compete be resurrected by the super-agency. However, there may be traffic volumes at which this latter situation could occur, since the super-agency does make use of its market power to increase profits. If one mode could at best incur a very small loss under competition then it would shut down. However, the additional power inherent in the super-agency could very well lead to its resurrection; however, to repeat, this does not occur in any case examined here.

6. Conclusions

We began this investigation with a number of questions relating to the organizational form and style of operation of Bay Area transit modes. On the basis of the rsults reported in Section 5 we may now summarize the answers obtained.

First, we wished to know if BART was capable of covering its operating costs, as it had originally promised. The answer, as is apparent from Tables 7 through 10, is that there exists a wide variety of circumstances in which this is indeed possible. One important case severely limiting the scope of this answer is when BART runs ten-car trains. Here BART is competitive only at the highest levels of travel volumes in the corridor in question. This suggests that the best strategy for BART is to run frequent, short trains.

Second, we observed that as presently operated, the BART system is imperfectly distinguished from the competing bus system (Table 1). The results here suggest that this is not an optimal policy for a BART system interested in covering costs. Given the revealed preferences of Bay Area commuters, and given that BART operates with small trains, Table 8 indicates that the fare could be raised considerably. This conclusion also holds for the case in which the bus system does not have to meet its operating costs.

Third, the question of the feeder system to BART arises. We observed that, as presently constituted, feeder service is provided by the bus company, having redesigned several of its routes to approach BART stations; and that an average fare of $.13 is charged (assuming that the average user makes a round trip daily, one leg of which is at full fare, $.25, the other free). The results presented here indicate that even when a feeder system is constrained to operate the same (large) vehicle types as the regular bus system, it always povides free feeder service, and provides better service in the sense of lower combined wait-and-walk time. The feeder service could presumably do even better by using smaller vehicles and thereby incurring lower operating costs. Thus, there may be considerable advantages to BART from integrating its linehaul and feeder operations.

Finally, there is the question of operating the transit systems as one through a super-agency. The results of Section 5.5 indicate that a transit monopoly, though it would lead to worsened service (and raised fares) on both modes, would not rely on a single mode only. Where "competition" would result in two modes, so would a single transit monopolist operate both modes.

These observations limit the scope of the redundancy questions raised in the introduction. There, to recall, the question arose as to whether it might not be better to keep both systems in operation when economic forces would, if allowed to operate freely, result in a single public mode of transportation being offered. However, the results of this analysis suggest that the question is relevant only when the BART system was constrained to run ten-car trains; and this, it was argued, was not the best BART policy. In the other cases, the transit market could support two competing public modes. Moreover, these modes would be able to cover their operating costs. Even when, for reasons of social policy, it was decided to ignore profitability considerations in the bus mode, a cost-covering BART operation could still be feasible.

NOTES

1. Webber (1976), pp. 1-2.

2. For example, in 1970, comparable distributions of income were:

Income	U.S.	U.S. Urban	San Francisco-Oakland SMSA
Over $10,000	15.3	16.5	22.5
Over $15,000	5.2	5.7	7.9
Over $25,000	1.5	1.6	2.2

Source: U.S. Survey of Population (1970).

3. See, for a survey of this history, Adler (1978).

4. Ibid.

5. Merewitz (1977).

6. AC Transit Annual Reports, 1962-1977.

7. Ibid.

8. BARTD Organic Act. Stat. 1957 c. 1056, p. 2309, Sec. 3. West's California Codes, Sec. 29038.

9. Pignataro and Falcocchio (1976).

10. Morlok and Viton (1980): Pignataro and Falcocchio, op. cit.

11. Morlok and Viton, op. cit.

12. BARTD Annual Report 1975/1976.

13. See, for this, Webber, op. cit.; Keeler, et al. (1975), Viton (1976).

14. BARTD Annual Report 1975/1976.

15. AC Transit Annual Report 1975/1976.

16. This apportionment is made by the Metropolitan Transportation Commission. Some government grants are also tied to property-tax collections.

17. Webber, op. cit., p. 6.

18. Adler, op. cit., p. 42.

19. See, for example, Intriligator (1971).

20. For a simple model of two duopolists providing substitutable products, see, recently, Dixit (1979). Dixit's analysis is for the linear case; while, as will be apparent, the problem to be discussed here is severely non-linear. It is interesting to note that in the Dixit model it is possible for a joint-profit-maximizer to produce none of one of the goods, while the duopolists produce positive quantities of both. This is the redundancy concept discussed earlier; as we shall see, it does not obtain in the BART/AC case.

21. See also Domencich and McFadden (1975), McFadden (1975).

22. Domencich and McFadden, *op. cit.*, p. 66.

23. *Ibid.*, p. 68.

24. *Ibid.*, Chapter 5.

25. McFadden and Talvitie (1977), pp. 148 *et seq.*

26. *Ibid.*, p. 148.

27. *Ibid.*, p. 139.

28. Domencich and McFadden, *op. cit.*, p. 124.

29. McFadden and Talvitie, *op. cit.*, p. 165.

30. *Ibid.*, p. 166.

31. *Ibid.*, pp. 169 *et seq.*

32. *Ibid.*, comparing pp. 172 and 150.

33. *Ibid.*, pp. 151 *et seq.*

34. *Ibid.*, p. 159.

35. For example, Keeler, *et al.*, (1975), Chapter 3, Section 3.1.

36. See, for example, Morlok (1978).

37. Three characteristics will be just the attributes revealed as important by the mode-choice model of Table 3.

38. See, for example, Intriligator, _op_. _cit_.

39. This model first appears in Pozdena (1975). It has received extensive use in the analysis of Bay Area transportation: see the references of note 13, _supra_.

40. Note that with facilities given, the modelling of transbay trips does not pose the same conceptual difficulties encountered in estimating the full resource costs of such trips (Viton, 1976).

41. The model is easily adapted to the analysis of other residential patterns.

42. There exists at least one such feeder service, from the University of California to the University Avenue, Berkeley, station. The difficulties of ensuring that such a service is used only for BART-bound travel are not discussed here. It is also assumed that the vehicle types run by the two bus operations are the same. For the costs of smaller buses, see, e.g., Fisher and Viton (1974).

43. Fisher and Viton, _op_. _cit_., pp. 18-22.

44. In fact, the seating capacity of BART cars is approximately 72. The procedure adopted here, due to Keeler, _et_ _al_., corrects for different amounts of space per rider provided by the two modes.

45. Reid (1977), p. 42.

46. Viton (1978a).

47. See note 42 <u>supra</u>.

48. Pozdena and Merewitz (1978); t-statistics calculated from Table I, p. 76.

49. Urban Transportation Factbook (1974), I-19.

50. Section 3.1, <u>supra</u>.

51. Thus there are built into the model assumptions about the nature and extent of itemized deductions, etc.

52. Census of Population (1970).

53. FHWA (1972). The actual probabilities ranged from .81 to 1.0 for the population sixteen years of age and older (the working population of interest here).

54. It is possible, though computationally expensive, to increase the fineness of the socioeconomic partition used here.

55. The numerical analysis used the ACDPAK algorithm, written by M.J. Best of the University of Waterloo. See Best (1975) and Best and Ritter (1974) for properties. Work performed on the DEC-10 computer at the Wharton School, University of Pennsylvania.

56. Merewitz (1975), p. 23.

57. In all the results presented here, answers are rounded as follows: headways (nearest minute); number of routes (nearest integer); fares (cents); mode share (nearest one percent). A single mode share is presented, averaged over all users, though the analysis is done in terms of the individual groups of travellers shown in Table 6. An equilibrium was attained when the mode choices of the individual groups were stationary to an error of one percent. The initial point for all problems was: Buses per hour, forty; Bus routes, ten; BART fare, $1.50; BART train size (when not constrained), two cars; Bus fare $1.50.

58. Keeler, _et_ _al_., _op_ _cit_., pp. 128-129.

59. That is, $.25 local fare when transferring from bus to BART and a free transfer in the other direction.

60. Of course, this assumes that AC Transit behaves in Cournot fashion. See _infra_ for BART possibilities when AC behaves otherwise.

61. The detail discussed below is based on the 5 x 6 matrix of mode-choice probabilities by income classes; of which only a summary appears in Table 8.

62. Why between 7,500 and 10,000 commuters per hour the bus service provided to satisfy the wait-plus-walk constraint changes so suddenly is not clear.

REFERENCES

Adler, S., "The Early History of Relations between the Alameda-Contra Costa Transit District and the SFBARTD," unpublished paper (1978).

Alameda-Contra Costa Transit District, Annual Reports (1960-1978).

Ash, R.B., Basic Probability Theory, New York: J. Wiley (1970).

Bay Area Rapid Transit District, Annual Reports (1969-1975).

Best, M.J., "A Feasible Conjugate Direction Method to Solve Linearly Constrained Minimization Problems," Journal of Optimization Theory and Applications (1975).

Best, M.J. and K. Ritter, "A Class of Accelerated Conjugate Direction Methods for Linearly Constrained Minimization Problems," Research Report CORR-74-12, University of Waterloo (1974).

Cluff, G.S., "Some Distributional Consequences of Alternative Pricing and Investment Rules in an Urban Transportation Market," Berkeley: University of California, Department of Economics, 1978.

Dixit, A.K., "A Model of Duopoly Suggesting a Theory of Entry Barriers," Bell Journal of Economics, Spring (1979).

Domencich, T. and D. McFadden, Urban Travel Demand: A Behavioral Analysis, Amsterdam: North-Holland Publishing Co. (1975).

Eads, G., "Competition in the Domestic Airline Industry," in A. Phillips, ed., Promoting Competition in Regulated Markets, Washington, DC: The Brookings Institution (1975).

Fisher, P.M.J. and P.A. Viton, The Full Costs of Urban Transport I, Berkeley: University of California, Institute of Urban and Regional Development, Monograph No. 19 (1974).

Hausman, J. and D.A. Wise, "A Conditional Probit Model for Qualitative Choice," Econometrica, (1978).

Hoachlander, E.G., "Bay Area Rapid Transit: Who Pays and Who Benefits?" Berkeley: University of California, Institute of Urban and Regional Development, Working Paper No. 267 (1976).

Intriligator, M.D., Mathematical Optimization and Economic Theory, Englewood Cliffs: Prentice-Hall (1971).

Keeler, T.E., et al., The Full Costs of Urban Transportation III, Berkeley: University of California, Institute of Urban and Regional Development, Monograph 21 (1975).

Keeler, T.E., "Domestic Trunk Airline Regulation," Draft (1977).

Lancaster, K., "Socially Optimal Product Differentiation," American Economic Review (1975).

Landau, M., "Redundancy, Rationality, and the Problem of Duplication and Overlap," Public Administration Review (1969).

Lansing, J.B. and G. Hendricks, "Auto Ownership and Residential Density," University of Michigan, Survey Research Center (1967).

McFadden, D., "The Measurement of Urban Travel Demand," Journal of Public Economics (1975).

McFadden, D., A. Talvitie, and Associates, Demand Model Estimation and Validation, Berkeley: University of California, Institute of Transportation Studies, Special Report UCS-ITS-SR-77-9 (1977).

Merewitz, L., The Full Costs of Urban Transport II, Berkeley: University of California, Institute of Urban and Regional Development, Monograph 20 (1975).

Merewitz, L., "On Measuring the Efficiency of Public Enterprises: Bus Operating Companies in the San Francisco Bay Area," Transportation (1977).

Morlok, E.K., Introduction to Transportation Engineering and Planning, New York: McGraw Hill (1978).

Morlok, E.K. and P.A. Viton, "Self-Sustaining Public Transportation Services," Transport Policy and Decision Making I:2 (1980).

Pignataro, L. and J. Falcocchio, "Survey of Ridership of Express Buses in the City of New York," Final Report, Polytechnic Institute of New York (1976).

Pozdena, R., "Choice of an Initial Fare Structure for BART," Berkeley: University of California, Institute of Urban and Regional Development, Working Paper 181 (1972).

Pozdena, R. and L. Merewitz, "Estimating Cost Functions for Rail Rapid Transit Properties," Transportation Research (1978).

Reid, F., "Disaggregated Supply Data Computation Procedures," Berkeley: Institute of Transportation Studies, Special Report UCB-ITS-SR-77-8 (1977).

Roberts, J. and H. Sonnenschein, "On the Foundations of the Theory of Monopolistic Competition," Econometrica (1977).

Schwann-1 Record Guide, March, 1977.

United States Department of Commerce, Current Housing Reports, H-170-75-39 (1979).

United States Department of Commerce, Statistical Abstract of the United States, (1978).

United States Federal Highway Administration, Highway Statistics (1972).

Urban Transportation Factbook, American Institute of Planners (1974).

Viton, P.A., "Notes on the Costs of Mass Transit in the Bay Area," Berkeley: University of California, Institute of Urban and Regional Development, Working Paper 269 (1976).

Viton, P.A., "Equilibrium of Supply and Demand in Urban Transportation," Berkeley: University of California, Workshop in Transportation Economics, Department of Economics, Working Paper No. SL-7706 (1977).

Viton, P.A., "Optimal Tolls on the Bay Bridge," Berkeley: University of California, Institute of Urban and Regional Development, Working Paper No. 297 (1978a).