Determining Cost-Effectiveness of Transit Systems

Paul N. Bay
Metropolitan Transit Authority
Houston, Texas

Cost-effectiveness is, by definition, a term with two components: "cost" and "effectiveness." Each component has both broad and narrow definitional possibilities. Further, the word "effectiveness" implies a goal or set of goals against which effectiveness is measured. To determine effectiveness, the intent, the effects to be achieved, must first be stated. Because transit goals are typically stated in broad terms, transit "effectiveness" is also likely to be broadly defined. Obviously, definitions are required if sense is to be made of the subject.

Presented in this paper is a range of possible definitions, with a brief discussion of the problems of measurement of each. There follow some simplified definitions and measurements that might be used in attacking the problem of deciding when light rail transit might be considered a cost-effective alternative.

COST

In considering the definition of "cost," and deciding what should and should not be included, each of the following categories is a possible candidate.

Initial Capital Cost

- Engineering,
- Right-of-way,
- Construction,
- Vehicles,
- Equipment, and
- Facilities.

Replacement, Renovation, and Upgrading Costs over an Extended Period of Time

- Vehicle replacement,
- Grade separations,
- Signals and train control improvements,
- Station additions or improvements, and
- Engineering and construction costs that are associated with them.

Operating Costs over an Extended Period of Time

These costs include such indirect costs as administration, overhead, taxes, and insurance.

Financing Costs

- Debt service on bonds or short-term borrowing,
- Bond counsel and other financing fees,
- Impact of restrictions resulting from bond protective covenants, and, possibly,
- Lost interest on investments or "opportunity costs."

System Cost Versus Marginal Costs

These are any or all of the previously mentioned costs applied systemwide and include all local feeder bus service, park-and-ride lots, and so forth versus the incremental cost of a given line, segment, or system portion over and above some lesser portion.

Total Cost Versus Local Cost

This is based on revenue sources: the portion of the total cost that is being paid by the implementing agency (local funds) or the amount being paid by the principal funding agency (federal funds).

Community Cost

This is total transportation system costs for all modes paid by public and private sources. This could be capital costs only or could include operating costs as well. This could provide a theoretically rigorous way of testing the thesis that transit expenditures could reduce highway and private automobile expenditures by taking the top off the peak travel periods.

Some have suggested that the community cost category might include other costs of any given transportation system mix, including construction impacts, levels of congestion, air quality, ambient noise levels, urban design and visual obtrusiveness, accident rates or other measures of safety, property values, job accessibility, economic development, freedom of mobility, social unity or division, neighborhood integrity and other "equality of life" factors, some of which are quite subjective and nonquantifiable.

This preliminary and not at all exhaustive list of possibilities addresses only the "cost" component of the term, "cost-effectiveness," but it does begin to illustrate the difficulties inherent in attempting to answer the seemingly simple question: "How much will this system cost?" Not only is the definition

of cost an important and complex matter, but the measurement of cost is difficult even when a definition has been agreed on. Even if all costs are measured in dollars, a decision has to be made about how to treat the time value of money, taking into consideration both inflation and interest rates. Years ago, Grant and Ireson (1) published the classic college textbook on engineering economy, but, once out of school, most engineers and planners have honored it only in the breach. Its principles have not been rigorously applied in either the highway engineering or the transit planning field. As a result, time value of money is usually not considered properly or consistently.

EFFECTIVENESS

To decide if a transit system is effective, it must be asked: "Effective in doing what?" There follow some of the categories that have been traditionally held as transit goals plus some that have not been considered as goals but perhaps should be.

Ridership Goals

- Total system ridership?
- Ridership on a particular route, guideway, or segment?
- Incremental ridership gains over some other option?
- Peak-hour ridership goals, daily riders, or annual riders per capita?

Corridor Capacity Increase

This is the ability to handle growth or reduce congestion, or both, in a given corridor or set of corridors.

Reduce Travel Time

Travel time reductions can be considered systemwide or in a particular corridor or set of corridors.

Increasing Connectivity or Accessibility by Transit

This could be for all trips, work trips, or some other set of trip purposes.

Reducing Environmental Impacts

Such impacts can be compared with those of a "highway improvements only" alternative, a "do nothing" alternative, or any other alternative.

Economic Development or Redevelopment

This could be anything from downtown revitalization to the ambitious goal of shaping or reshaping a region or a portion of a region to make it more generally sellable through tourism, new jobs, new kinds of industry, and so forth.

Solving Political Problems

These goals can range from reallocating or redistributing wealth, providing for greater social equity, equalizing services, and dealing with tax inequities, to building monuments or helping powerful interests.

Reducing Total Transportation Costs

This goal would be to reduce the total cost to the community of providing a certain level of mobility or accessibility for all or most citizens. As in the "cost" side of the equation, this goal could be limited to reduced capital expenditures by the public for the combination of highway and transit facilities or it could be broadened to include both public operating and private transportation costs--automobile amortization, insurance, fuel, and maintenance.

GOAL SETTING

The conclusion inevitably reached from the preceding is that it is literally impossible to prove when a transit system is cost-effective unless the set of goals to be achieved has been defined in as precise and quantitative a way as possible and estimates have been made of the cost of achieving those goals under each of several different transit-transportation scenarios.

The major problem with this is that goals have not been well defined. Goals, as defined, have tended to be general and nonspecific (i.e., "improve environmental quality," "reduce congestion," "enhance urban mobility," and "develop downtown"). Alternatively, goals have been defined in a manner that defeats the purpose and thereby loses credibility and public understanding. The primary example of this is an emphasis on line ridership as the only publicly stated goal. Ridership is important--probably the single most important goal that is both easily measurable and central to other goals. However, if systemwide goals are not understood, knee-jerk reactions to early ridership figures can be highly misleading.

CALCULATING COSTS

Not only have goals not been defined well, but results of calculations of costs in the transit cost-effectiveness analyses done nationwide have been wildly at variance with one another. Clearly, there is no agreement among even knowledgeable academics, consultants, and partitioners about what items should be included in the answer to the question: "How much did that transit system cost?"

Most agree that initial line capital costs should be included but differ on whether and how to include line operating costs, local bus feeder service costs, financing costs, or total transit system costs. The time value of money is also treated differently, as was mentioned previously.

DEFINITION AND EVALUATION OF ALTERNATIVES

The problems of goal setting and cost determination are difficult, but what really leads to never-never land is the task of defining and costing alternatives to any actual or proposed transit system that are capable of achieving the same set of goals.

If alternatives are defined and costed, theoretically it can be determined whether the transit system under consideration achieves the goals at less cost or at greater cost than do the alternatives. Without comparing a specific transit system to such alternatives, only the goals the system under consideration

achieves and at what cost can be stated. Whether or not that cost is "cost-effective" is a value judgment determined by what any individual or group is willing to pay for achieving those goals.

The difficulty comes in addressing "what if's": "What if we had built a busway instead of a light rail line?" "What if we do nothing but stick with the status quo?" "What if we just widen the freeway?" Deciding "what if" means guessing the effects that ensue when only one variable is changed in a highly complex mix of hundreds of dependent and interdependent variables that make up an urban social and physical setting. This was nowhere better illustrated than in the fierce--and ultimately futile--arguments that raged among the planners, economists, engineers, and political scientists hired to do the Bay Area Rapid Transit (BART) impact study over the problem of defining the so-called "No-BART Alternative." If BART had not been built, would a new bridge across San Francisco Bay have been built instead? Would AC Transit's bus system have been further expanded? Would the San Francisco and East Bay freeways have been wider? Would there have been fewer high-rise office buildings in downtown San Francisco? Would more or fewer people be living in the suburbs? And what would have been the costs and impacts of any or all of those other things? To this day, some well-known academics at the University of California, Berkeley, who were involved in those debates periodically renew their luster (or notoriety) by delivering themselves of pontifical opinions about BART. These opinions gain them attention but have no more relevance to the real world than do medieval monks' arguments about the number of angels that can be accommodated on the head of a pin.

WHAT THEN?

It could be concluded from the foregoing that a cost-effectiveness determination is hopeless and that the effort should be given up, but that would not be the author's intention or viewpoint. There is hope, in spite of the difficulties, that analysis can help overall understanding and improve decision making. Common sense and consistency, more than rigorous and theoretically pure conclusions, are needed. A better job can be done than has been done in the past. Two seemingly contradictory recommendations follow.

Broaden the Analysis Base

In this author's judgment, much of the previous work in evaluating transit systems has been off the mark because it has been too limited in terms of both costs considered and effects produced. Transit alone has been examined, not the total transportation system of which it is a part. (Some would go even further and look at all the land use, environmental, social, physical, and economic systems with which the transportation system interacts, but that raises too many of the difficulties described in the BART impact studies.)

There are real-world definable trade-offs between transit costs and highway costs for both public entities and private individuals. These have not traditionally been viewed in terms of systemwide transportation. However, to do justice to this subject, a long-term view must be taken--probably 25 years as a minimum--and more research is needed. A fascinating research project that should be undertaken would be to chart the total public and private expenditures for transportation since the end of World War II in a range of metropolitan areas--some of which have chosen higher transit expenditures and some of which have put all their eggs in the highway basket. Toronto, Washington, San Francisco-Oakland, and Atlanta might be chosen for one category, and Houston, Los Angeles, Denver, and Seattle or Dallas might be chosen for the second category.

Such a research project would consider all expenditures, including construction, maintenance, upgrading, and repair, for freeways, arterial streets, and bridges, and the transportation equipment and facilities and manpower used by state highway departments, local governments, and private developers. It would include all transit costs in a similar way: bus purchase and replacement, garages, rail lines, operating and maintenance costs--the works. Then it would use fare-box revenue, gasoline sales data, vehicle registration figures, insurance industry records, and "Hertz-type" automobile operating cost data to calculate private transportation costs, taking care to avoid double counting. Some judgments would have to be made about including such things as commercial parking lot construction and local residential streets, but the decisions would probably be less important than consistency among all the cities.

Then, when the cost side of the ledger for all these metropolitan areas has been examined, "effectiveness" or goal attributes could be looked at: congestion levels, travel times, mobility levels, job choices within 30-min travel time, and so forth. Although such as analysis might still beg the question of which type of city is "best," it would clearly show the total transportation costs and the results over an extended period of time. It would not be surprising to find that the cities that spent higher levels of money on transit actually spent less in total on all transportation and achieved comparable levels of personal mobility.

Narrow the Analysis Base

In the absence of long-term research information as just proposed, decisions still have to be made. To do this, a narrowing of the analysis is necessary, and this is forgivable if consistency is maintained from location to location. The cost-effectiveness criteria proposed by UMTA in the new alternatives analysis requirements are a good starting point for such a short-hand approach. Some transit planners will dispute this judgment, and, of course, improvements are possible, as suggested next.

The UMTA criteria focus, first, on segment capital costs and marginal operating costs and, second, on marginal ridership increases. Those are measurable, and they fit the reasonableness test of the average person. They are understandable. As such, they constitute a good start in a simplified approach to making judgments about cost-effectiveness.

The area that is left out of the UMTA analysis is the marginal cost impact of the transit investment on total transportation system costs. Such an analysis is possible in a simplified form as well as in a long-term and comprehensive form. For example, if construction of a light rail line will eliminate the need to add two freeway lanes in each direction in the same corridor to handle projected peak-hour demands, and if it can reasonably be shown that the cost of building those added lanes is a certain amount, that amount should be included in the UMTA alternatives analysis--not just transit alternatives versus other transit alternatives.

If the UMTA cost-effectiveness criteria were used and the marginal cost impact on highway improvement requirements were added to those criteria, the results might be closer to the mark.

SUMMARY AND CONCLUSIONS

An examination of the literature makes it clear why there are so many arguments about transit's cost-effectiveness. Lack of clarity and consistency in definitions, measurements, and methodology has characterized the whole field for 20 years. The author recommends three things to reduce the present ambiguities:

1. Transit cannot be examined in isolation, but only as part of the total transportation system for any community--costs and effects must be broadened to include the highway and automobile part of the system. However, this broadening should not try to also include social, environmental, and economic costs and effects in a rigorous way. Such factors can be examined in a subjective, judgmental manner, but that should be separate from the quantitative analysis of the transportation system.
2. To do a better job of understanding the total costs and effects of alternative transportation systems, some broader, long-term research is badly needed.
3. In the shorter term, the UMTA cost-effectiveness criteria represent a good start toward greater consistency although they lack the broad base that research might provide. However, the UMTA cost-effectiveness criteria should be modified to permit inclusion of related marginal highway cost impacts in a manner consistent with the treatment of marginal transit cost impacts.

REFERENCE

1. E.L. Grant and W. Ireson. Principles of Engineering Economy, 5th ed. Ronald Press Company, New York, 1970.