Mike J.

Lessons from SMD Experience with Approaches to the Evaluation of Fare Changes

Lawrence Doxsey
U. S. Department of Transportation
Transportation Systems Center
Kendall Square
Cambridge, MA 02142

prepared for

Urban Mass Transportation Administration
Office of Technical Assistance Programs
Service and Methods Demonstrations Division

Introduction

Over the past several years UMTA's Service and Methods Demonstration Program (SMD) has undertaken a large number of studies of the effects of fare changes, both increases and decreases. Some of these studies have been large scale efforts directed at identifying not only aggregate ridership and revenue impacts of fare changes but also at improving understanding of the complexities of individual travel behavior. Because of their scope they have required the simultaneous application of several methods. Others have been directed strictly at aggregate revenue and ridership forecasting for medium term planning purposes. In the course of pursuing this program of evaluation much has been learned about the relative advantages and disadvantages of alternative evaluati n approaches.

In general, approaches to evaluating the effects of fare changes, most commonly fare increases, can be assigned to one of three categories. The three are time series analysis, comparisons of independent cross sections, and panel studies. Through SMD sponsored evaluations experience has been gained with each of these. The methods are distinguished not only by the data used and the complexity of analysis required to implement them, but also by the kinds of insight they offer and by the costs of their implementations. Because different information is provided by each of them, they are not interchangeable in application and care must be taken to match tools to purposes. In what follows each of these approaches will be discussed in turn.

Perhaps the most emphatic conclusion from the SMD experience is that there can not be found a heirarchy of methods under which increasingly elaborate and expensive methods fully serve the purposes of simpler methods as well as adding further complexity. Rather, each method can be allied with particular applications and achievement of multiple purposes generally requires multiple methods, although with a broad approach it is true that some symbiosis is commonly evident.

Time Series Methods

For simple measurement of the relationship between fare and ridership, and hence that between fare and revenue, estimates based on a time series of aggregate measurements dominate those of other methods. In most applications two to three years of monthly data is sufficient to accurately estimate the necessary parameters. The most powerful estimation tools for this purpose are generalizations on Box-Jenkins techniques which allow inclusion of some causal variables. However, even ordinary least squares regression can provide quite good results.

With monthly observations, data requirements commonly include total ridership, the "real" level of fare obtained by deflating the cash fare by an appropriate price index, the total employment level for the system's service area, and the number of working days in each month. In periods for which energy prices have been volatile, inclusion of the "real" price of gasoline is warranted. Most of these data are readily available so that time series methods are quite inexpensive to apply. This is in particular constrast with the independent cross section and panel methods which tend to be costly in both data collection and data analysis.

Because it measures the relationships between fare and and some aggregation of ridership, and thus between fare and revenue, the time series model provides an effective planning tool for predicting the ridership and revenue consequences of alternative fare levels. For the purpose of forecasting ridership and revenue impacts, time series methods are both more precise and less costly than either of the other two methods considered below.

Furthermore, time series techniques are fully applicable to evaluation by service type or even by route, provided the data used are disaggregated to the appropriate level. Thus not only can they be used to generate aggregate forecasts of ridership and revenue resulting from alternatively considered fare levels, but they can do the same thing for disaggregations on categories such as service type. For example, to measure fare elasticity differences between express and local service, it is necessary only to have ridership and fare data by these two service types.

The prime drawback of time series methods is that they provide no insight into the relative impacts of fare changes on different groups of riders, such as on low income versus high income riders. Thus by using a time series method alone it is impossible to distinguish whether a contemplated fare change impacts one group more heavily than another. It is in this area of disaggregate information that the other two methods compare favorably.

Independent Cross Sections

When conducted on a before/after basis relative to a fare change, a sequence of two or more separate surveys of transit riders allows identification of the changes in ridership composition coincident with the fare change. Thus, for example, a pair of independent cross sections bracketing a fare increase could be used to identify a decline in the share of non-work trips or an increase in the share of trips taken by high income riders. While no causal attribution could strictly be made, in most situations there would be strong heuristic support for attributing the greatest portion of any changes in ridership to the change in fare.

Because the focus of this approach is entirely on transit riders, on-board surveys are an appropriate mechanism for data collection. With self-administered surveys this means that the method of comparing independent cross sections will be only moderately expensive to implement. In this context, it is worth noting that a crucial consideration with on-board surveys, although one frequently overlooked, is that its sampling procedures give known selection probabilities to individual trips but not to individual users. Through tabulation of on-board survey results one obtains pictures of the distributions of trips made by users in various socio-economic categories but one does not obtain pictures of the distributions of the socio-economic characteristics of the population of users. While statistical corrections are available, on-board surveys remain best suited to profiling trips made on a transit system, not the system's users.

Note that the method of independent cross sections does not provide a replacement for a time series method. It can be used neither to accurately measure aggregate ridership and revenue impacts nor to forecast the ridership and revenue impacts of comtemplated fare changes. Instead, it complements a time series approach by providing insight on the distributional consequences of a fare change through identifying changes in the composition of ridership coincident with the fare change.

Panels

By definition, a panel involves a sequence of at least two observations on a single group of individuals. In comparison with independent cross sections, by assuring that before/after observations are made on the same individuals, panels allow disaggregate measurement of changes in travel (or other) behavior. Any of a variety of circumstances of the individual can be associated with the observed behavioral change and quite complex travel behavior relationships can be studied. Panel designs readily lend themselves to estimation of causal relationships at a disaggregate level. Thus a well designed panel can separate the effects of changes in job location from those of a fare increase as factors influencing changes in individual transit usage. Neither time series methods nor independent cross sections can even approximate this causal separation. Because they include as exogenous variables both intertemporal changes in various circumstances which may affect individual travel behavior and data on individual characteristics, panel designs permit identification of the influence on behavior of factors other than fare. At the individual level, then, the influence of fare is separated from the influences of other forces. However, isolation of these other effects may be of little interest for purpose of transit planning and operation, relegating panel approaches to rather more academic and less operational studies of fare changes.

The two major disadvantages of panel methods are that they are expensive for both data collection and analysis and that it is, at best, difficult to implement them in ways which allow aggregation from the panel's sample to total population. Many of the actions which can be taken to improve a panel's suitability to aggregation also lead to very great increases in expense so that attempts to relieve the second disadvantage tend the aggravate the first.

An example may serve to illustrate both the strengths and limitations of the panel method. For evaluation of a fare increase, a deceptively attractive panel design is to conduct an on-board survey and to solicit either addresses or telephone numbers from survey respondents so that they can be subsequently recontacted either through the mail or by telephone. The combination of initial on-board survey and the subsequent follow-up provides before/after measurement on each of a set of individuals and thus constitutes a panel design.

One imposing barrier to the ultimate utilization of this approach for estimating aggregate relationships is the likelihood that a substantial share of the target population will refuse participation. This refusal has three components. First is the refusal to participate in the on-board survey. SMD experience suggests that response rates to self-administered on-board surveys range roughly from 25 percent to 75 percent but cluster in the 40 to 60 percent range. On-board survey refusal is a problem which this particular panel design has in common with the method of comparing independent cross sections. The second component of refusal is the share of participants in the on-board survey who do not provide the requested address or telephone number. The third component is the share of those who provide the requested recontact information but who either cannot be reached or who refuse participation in the follow-up survey. Note that these three components are multiplicative in their effects. Thus, for example, with a 50 percent response to an on-board survey, with 30 percent of the on-board respondents refusing to provide the information necessary to recontact them, and with 30 percent of those who provide the information needed to be recontacted ultimately either refusing to participate in the follow-up survey or not able to be recontacted, the effective refusal to panel participation is somewhat above 75 percent. To use the panel for generalizations about the population implies either assuming that the minority of the population potentially in the panel are indistinguishable from the majority, or that the differences can be somehow identified. The former is an indefensibly bold assumption and the latter can never be more than approximately the case. In general one is left simply not knowing the relationship between the panel and the population and thus left without a means to use the former to make inferences on the latter.

A second limitation with the illustrated design is that because the initial contact is with an on-board survey, it shares with all on-board surveys the characteristic that while each trip has a known probability of being drawn, each person's relative probability of inclusion is proportional to his/her boarding frequency. While this limitation can be statistically overcome in a number of analysis methods, doing so adds complexity to the analysis process. Furthermore, overcoming the problem in analysis of the panel does not remove it as a barrier to aggregation from the panel to the total population.

Notice that the design includes people who cease use of transit due to changes in external factors such as employment location but excludes those who begin use due to changes in external factors. Unless there are joint influences from the simultaneous fare change and changes in external factors, this assymetry will not hinder the design's ability to isolate the separate influence of fare on individual usage.

The example of this particular design illustates some of the difficulties in using panel methods. The limitations evident in the example can be circumvented only incompletely and then only at considerable expense. For instance, interviewer administration of the on-board survey can reduce the first stage of refusal but requires considerable commitment of labor. At an even greater effort level, initial contact through a household survey can both substantially improve the overall response rate and eliminate the problems of on-board survey sampling idiosyncracies. For some applications these expenses may be justified. However, even with substantial effort aggregations based on panel results will only approach the accuracy of time series analysis. While the panel will provide a far richer insight into the complexities of individual behavior lying behind the net effects reflected in the aggregate, for many purposes much of this richness has no bearing on the relationship between fare and ridership.

For example, in using a model built on panel data to forecast the effect of an alternative fare level, one would reasonably hold constant all independent variables but for fare, project the response of each panel member, and aggregate across the panel to estimate the total impact. But this procedure does not use any of the information in the panel other than the relationship between fare and ridership. If there are trends in the net changes of other factors which influence ridership, the panel based forecasting approach will not reflect them unless the trends can be assigned as changes for individual panel members. Thus the approach will provide the same answer with or without changes in the other factors which have been identified as influencing ridership. In comparison, a time series approach will reflect the influence of trends continuing from the the period on which a model is estimated into one for which a forecast is made. It will not, of course, reflect any change in trend occurring between the two periods.

Summary

SMD experience with approaches to evaluation of fare changes suggests the following conclusions. Time series approaches provide accurate and inexpensive estimates of the relationship between fare and aggregated ridership or revenue. For purposes such as setting differential fares for several service categories so as to maximize ridership while meeting a total revenue constraint, time series methods dominate both comparisons of independent cross sections and panel methods.

Comparisons of independent cross sections developed through before/after on-boards provide 'snapshots' of transit usage which enable identification of overall changes in patterns of usage coincident with a change in fare. This can be useful for assessing the distributional consequences of changing fare. However, the method does not lend itself to causal interpretation so that any attribution to a fare change of observed changes in usage patterns must be tentative.

Coupled with appropriate analysis, the panel approach identifies individual response to fare increase while controlling for the influence of individual specific changes in exogenous characteristics such as employment status and auto ownership. Furthermore, it allows identification of differences in these relatonships among different population segments. Thus relative to comparisons of cross sections, panel methods have the potential for isolating causal relationships between fare and ridership. On the other hand, use of the panel approach provides very weak identification of the incidences of various characteristics across the population so that the relationships identified at the individual level can not be scaled up to discern impacts at the full population level. Additionally, even when done with minimal embellishments, panel studies require expensive data collection and expensive analysis.