P.I. Peter Stopher, Ph..D., and Stephen P. Greaves
Louisiana State University

LTRC Contact:
Art Rogers, P.E.
Phone (225) 767-9166

INTRODUCTION

Metropolitan areas with populations of over 50,000 are required to conduct transportation planning on a continuing basis to maintain a transportation plan that qualifies for federal funding. Data to support such planning are expensive to collect and can become outdated. This research was initiated to assist metropolitan planning authorities in Louisiana to acquire the necessary transportation planning data as cost-effectively as possible.

OBJECTIVE

The objectives of this study are:

1. To develop and validate a methodology for MPO's to synthesize household travel survey data using local sociodemographic characteristics in conjunction with a national source of simulated travel data.
2. To evaluate whether travel-demand models estimated with these synthetic data perform better than models or statistics developed in other contexts.
3. To evaluate the transferability of the approach. Here, the procedures will be run for two other urban areas with different socio-demographic characteristics to Baton Rouge - Baton Rouge is the test site for meeting the first two objectives.

RESEARCH APPROACH

The methodology followed to develop and evaluate the concept of creating synthetic household travel survey data was broken into three phases. In Phase I, procedures were tested and evaluated for simulating the survey data set. In Phase II, models calibrated with these synthetic data were compared to current models and models calibrated using an actual travel survey data set. In Phase III, the approach was extended to two other urban areas of somewhat different characteristics to Baton Rouge to evaluate the transferability of the procedures.

CONCLUSIONS

This research proposes a new approach designed to assist practitioners in preparing travel-forecasts for their region. Rather than focusing on aggregate relationships and models per se, the idea is to synthetically derive the data that drives these models by combining local sociodemographic information with simulated travel data. The appeal of the approach lies in its low-cost, relative ease of use, the fact that the data required are freely available – NPTS (Nationwide Personal Transportation Survey) and PUMS90 (Public use Micro-Data Sample) – and the potential advantages it offers over the use of borrowed models that incorporate no local demographic element directly in their estimation process. In addition, it provides regions with the

capability to specify and estimate their own models rather than being tied to structures employed in transferred relationships. Finally, the approach offers a flexibility that will enable regions to regularly update their travel data base and modeling efforts as new demographic and travel data sources become available, such as the 2000 waves of the PUMS and the NPTS (now the NHTS).

The following issues affected the simulation and must be brought to the reader's attention. First, the approach is based on the premise that people of similar demographic characteristics living in similar environments display similar travel behavior. If this is so, then the problem should simply be one of correctly defining these characteristics. However, the indicators for the segmentation schemes suggest a limited ability to explain the variation in certain types of trips (particularly to shop and other) with the characteristics available. These differences must be attributable to other factors that are not captured in conventional demographic measures such as health status, personality, and whether the individual is planning for a particular event. A further complication is added by how the dependent measures are defined.

A second issue that affected the simulation was the use of the NPTS as the source for the simulated travel data. While it proved a particularly "clean" data set, the collection methodology was somewhat different than the typical travel survey providing data for a national cross-section for every day of the year. More significantly, however, use of this data source assumed that nationally-derived relationships were maintained at a local level. This appears to be partially true and suggests that some local information must be employed to ensure the synthetic data display sensitivity to a particular locale. Another possibility might be to apply the same logic as model transfer and use a travel survey from a region of similar characteristics (if available) or an amalgam of several travel surveys

as the source of the simulated travel data.

Another issue that affected the simulation was the decision to work at the household level. For trip rates, it is arguably more appropriate to work at a household level because of the effect of interactions among household members. However, for mode, departure time, and trip length, it could be argued one should work at an individual level, although one does still need information on household-level variables such as the number of automobiles. Comparisons between the simulation results using households and individuals is recommended as a potential future research topic. Clearly, while the concept of synthetic travel survey data is a seemingly valid one, further testing/evaluation is needed. The approaches presented here were intended for application using the 2000 waves of the PUMS and the NPTS. However, in the future, another possibility for updating the demographic database will come from the American Community Survey (ACS). The ACS will provide PUMS-like data for three million households per year sampled from across the nation. This could provide a means to update the database annually rather than every ten years.