1 INTRODUCTION

From different reviews about the explanatory factors for the dynamics of public transport patronage available in the literature [Webster and Bly, 1980; Goodwin, 1992; Nijkamp and Pepping, 1998; Dargay and Emly, 1999; etc.], it appears that time-series analysis gives interesting results on fare elasticity, while cross-sections show differences due to the various components of the quality of service and of the competition between transport modes. Thus, useful methodologies depend on the data available and on the effects considered as most important, including economic growth which has not been taken into account in the studies listed above. Another factor is the structure of the population: urban sprawl and the declining proportion of young people have a negative impact on the use of public transport, which has to be introduced in the models. Finally, it is important to consider a multi-modal framework: the private car is the main competitor of public transport.

In order to tackle these problems, a mainly demographic approach has been developed at INRETS for forecasting car ownership [Madre et al., 1994] and urban mobility [Bussière et al., 1995]. It takes into account structural factors in a multi-modal context, but does not allow to elaborate scenarios with different hypothesis on economic growth, nor with different price and supply policies. The two econometric approaches presented here aim to integrate these demographic and economic aspects. Let us first describe the data on which demand forecasts for public transport in the metropolitan area of Paris can be based. Then two methodologies will be presented with
different estimates of long term elasticities. Finally, considering the accuracy of these parameters (confidence intervals) the uncertainty of forecasts will be discussed.

2 THE TIME SERIES DATA AND THE MODEL

Both the accuracy of the estimates for the parameters of the models and the number of variables which can be introduced depends on the quantity of data analysed. Since too few data are available before the 70's, a single equation adjusted on annual time-series does not allow to consider as many explanatory factors as would be needed. Considering competing traffics, Paris is the only French metropolitan area where statistics on road traffic exist for a long enough period to allow econometric estimates. It is available on a monthly basis since 1980, as well as the traffic of public transports (except for private busses operating in the outer suburbs, which represent about 6% of total PT traffic). Thus, a time-series approach with a system of two equations (road and PT) has been implemented on the period 1980-96.

Concerning public transport, we have also collected annual time-series from 1975 to 1995 for 62 French urban areas including Paris. Thus, panel econometrics methods can be used for the calculation of elasticities. Traffics are measured in terms of vehicle*km (for road in Paris region), passenger*km (for PT in Paris region) or number of trips (PT for the 62 cities).

Explanatory factors are:

Income (from fiscal sources);

For public transport,

1. quantity of supply (in seats*km), but also quality of service (frequency and network density);

2. for road, the toll parking space inside Paris and the length of motorways in the suburbs;

3. prices for PT fares and fuel (competition with car use).

4. The indicator of population structure is a combination of:

   a. Numbers of individuals for a segmentation of the population relevant for public transport analysis [Madre and Boulahbal, 1999]: i.e. age and gender, the number of cars in the household and home location within the public transport area (central area or suburbs);
b. "mobility" coefficients for public transport use (number of trips per person and per week).

The indicator of population structure is the sum of these population categories weighted by their mobility coefficients. The mobility coefficients are kept constant while the proportion of each category of the population varies over time.

By introducing this single variable, we take into account inter-related factors, which have mainly a negative impact on the use of public transport:

1. more cars available,

2. more households living in the suburbs,

3. less young people.

These factors are partially correlated: for instance, the growth of car ownership is stimulated by urban sprawl. Thus, introducing different variables to measure the impact of these factors would have meant collinearity problems and loss of degree of freedom.

The estimated model is a system of two equations which explain both traffics of public transport and road traffics using the previous explanatory variables. This system is estimated with the Generalized Method of Moments. Results are robust and estimated traffics are very close to observed traffics, especially for the seasonal peaks. This estimated system gives us a good aggregated view of the monthly demand for road and PT but implicitly considers a uniform scheme for the demand for public transport in Paris region.

3 A PANEL OF 62 CITIES

The benefits of using panel data are numerous. They provide more information than either cross-section or time-series data, thus allowing more complex behavioural model specifications, while providing more reliable parameter estimates.

Panels permit the investigation of the dynamics of adjustment, which is impossible with cross-section data, while maintaining individual variation, which is lost in aggregate time-series data. Finally, the use of panels allows controlling for individual and temporal heterogeneity, which would otherwise lead to biased estimates. Most commonly in panel data analysis, either a fixed or random effect model is used to account for heterogeneity. In the former, the heterogeneity between cross-section units and/or time periods is specified by individual and/or time specific
intercept terms. In the latter, the heterogeneity is captured by individual and/or time specific error components. Both of these specifications assume that the coefficients of the explanatory variables included in the model are the same in all cities and over time.

These data do not support this hypothesis; instead, the coefficients appear to be different among cross-section units or to vary over time. In a traditional regression model, it is statistically impossible to allow for both individual and time-varying coefficients, as it would require the estimation of more coefficients than there are observations. Even the possibility of estimating either individual or time-varying coefficients is limited, as it would require either a long time period compared to the number of cross-section units, or vice versa.

In order to circumvent this problem, instead of treating the regression coefficients as fixed variables, they can be viewed as random variables with a probability distribution. This 'random-coefficients' approach greatly reduces the number of parameters to be estimated, while still allowing the coefficients to differ amongst cross-section units and/or over time. In this study, a Bayesian method is used, in which the parameter estimates are a weighted average of pooled regression and individual time-series regression, so that each individual regression is 'shrunk' to tend to the pooled regression estimates.

The use of these shrinkage estimators allows the estimated coefficients to vary across areas, while not sacrificing the efficiency of the estimates. Two different functional specifications are estimated: a log-log model, which constrains all elasticities to be constant over time and for all values of PT use and the explanatory variables; and a semi-log model, in which the fare variable is specified in levels, but all other variables as logs, so that the fare elasticity increases with the fare level, thus varying over time.

Surprisingly enough, the elasticities for Paris are in the middle of their distribution, although Paris is by far the largest urban area in France. For Paris, the estimated short-run elasticity of PT demand relative to fares is –32%, to seats-km per capita is 33%, to the generalized frequency is 10%, to the structural mobility is 28%, to the oil price is 5.7% and to the density is 5.4%.

4. Uncertainty of forecasts

At first, the elasticities derived from thes two approches seem quite different: time-series show a higher sensitivity to income and a lower to prices and supply variables. The elasticity to the indicator of population structure is not significantly different from unity when it is calculated on time-series, which is not the case in the panel approach (around 0.4 on the long run). Although very different, these estimates are not inconsistent because the confidence interval calculated from time-series is very wide.

Now, we have to investigate the different sources of distorsion between these estimates:
1. the time unit (a month for the time-series approach, a year for the panel approach),
2. the period covered by observations (1980-96 for time-series, 1975-95 for the panel); we have already found quite different results from the panel approach on the period 1987-95 compared to 1975-95 [Bresson et al., 2000];
3. the set of explanatory variables is not exactly the same in both studies; anyway, the explicit impact of car traffic is specific to the time-series approach; the competition with car use is seen only through car ownership (an important component of the indicator of population structure) and fuel price for the panel approach.
4. Finally, we will compare confidence intervals of public transport traffic forecasts obtained for both approaches.

REFERENCES

Armoogum J., Bussière Y., Gallez C., Girard C. and Madre J-L. (1994), Longitudinal approach to motorization : long term dynamics in three urban regions” International Conference on Travel Behaviour (IATBR), Valle Nevado, Chile.


