## Approaching a Dynamic Urban Transit Demand Model for Sydney

**Ilker Altinoglu**
*Institute of Transport Studies*
*University of Sydney*

**Nariida Smith**
*Senior Lecturer*
*Institute of Transport Studies*
*University of Sydney*

---

Abstract:

Urban Australia has seen a continuing movement away from public transit. In 1988 over 95% of all passenger kilometres in the Sydney Metropolitan area were by car and truck. There is now a growing recognition of the costs of increased automobile use both locally in terms of congestion, pollution and accident costs and globally from vehicle greenhouse gas emissions. Strategies to stem the trend must come from a clear understanding of all the factors effecting demand in the long and short term at the micro-economic level

This paper discusses the approach to be used in development of a dynamic model system for transit demand in Sydney using travel survey data from 1971, 1981 and 1991 Use of data spanning 20 years means that the effect of land use on transit demand can be examined The model system will aim to allow analysis of questions regarding public vs private transport not as "either/or" but rather in terms of providing the most appropriate mode for the context

---

Contact Author:

Dr Nariida Smith
Institute of Transport Studies
Graduate School of Business
The University of Sydney NSW 2006

Telephone: (02) 550 8686 Fax: (02) 550 4013

---

## Introduction

Urban Australia has seen a continuing movement away from public transit towards more use of the private motor vehicle. In 1988 over 95% of all passenger kilometres in the Sydney Metropolitan Area were by car and truck. Road infrastructure investment has not kept up with the growth in automobile ownership and use resulting in increased costs to the community in terms of accident costs and congestion costs. In addition to these local costs there has been increasing recognition of the global costs of vehicle emissions contributing to global warming via "the greenhouse effect".

Use of public transport is often proposed as a solution for decreasing vehicle use. Some proposals are of the "everyone else should use public transport" variety. Others uphold the commonly held belief that low density urban development in cities in Australia and the USA makes them less suited to public transport than the medium to high density European cities. However evidence such as Newman and Kenworthy (1989) suggests there is considerable difference in public transit use among seemingly similar urban environments. Few proposals address the full complexities of the issues involved.

People choose the mode of transport which gives them maximum utility subject to money and time budget constraints. The factors which determine this utility are a combination of the attributes of the form of transport, and the characteristics of each person including their environment. Any strategies to increase demand for public transport should stem from a clear understanding of these interactions both in space and time. Time must be considered since there is often a time lag between changes of circumstance and changes in behaviour. Now is a perfect time to develop such an understanding. Extensive Sydney wide travel behaviour data for 1991 stemming from a survey being conducted by the NSW Transport Studies Group is soon to be available to complement similar surveys carried out in 1971 and 1981. This will provide an opportunity to study changes in behaviour over time. This paper discusses the approach to be used in development of a dynamic model system for transit demand using the Sydney travel survey data spanning twenty years.

## Historical perspective

There has been a growth of interest in the dynamic modelling of travel behaviour since the 1970s. Several researchers considered the importance of change as a dynamic process. Goodwin and Mogridge (1981) and Clarke et al. (1982) discussed problems that could prevent static analyses from uncovering causal relationships and suggested dynamic approaches to the forecasting of travel behaviour.

Dynamic behavioural modelling has been developed in the research of Hensher and Wrigley (1986), Hensher (1988), Hensher and Smith (1990) and Hensher et al. (1992) in the Dimensions of Automobile Demand Project (1981-1991). Four waves of panel data collected in the Sydney Metropolitan Area during the period 1981-1985 were used to jointly estimate discrete choices of type and number of household vehicles and the continuous choice of their amount of use.

Increased availability of longitudinal data collection especially panel data sets enabled an increasing number of dynamic analyses of travel behaviour. The National Mobility Panel of the Netherlands was instituted in 1984 to study the changes in the mobility of the Dutch population. A description of this panel was given in Golob et al. (1986).

Goodwin (1989) used a dynamic analysis model in the context of the Dutch panel data available at the time to demonstrate that the changes or transitions in the circumstances of the individuals and households such as in their life-cycle, employment status, car ownership or income had effects different from those expected from a single-cross section analysis.

A series of research papers were published, especially by Golob and co-researchers, to incorporate dynamic structural modelling using the Dutch Mobility Panel (1984-1987). Golob et al. (1986) focused on identifying patterns of change in the use of various modes of transport, using categorical variables and log-linear models. The dichotomous use/non-use variables were used to generate turnover tables for each mode and different population segments. The data analysed was collected in three panel waves six months apart in 1984 and 1985.

The same data and demand for different transportation modes was studied by Golob and Meurs (1987) by using a simultaneous equation model. The temporal trends and shifts of demand among modes were analysed in terms of three different types of relationships: contemporaneous links among mode demands at the same point in time; temporal links among the demand levels for the same mode at different points in time; and cross-lagged links between one mode at one point in time and another mode at a later point in time. The structural relationships in the form of linear simultaneous equations were estimated using the LISREL program (Joreskog and Sorbom 1984). Golob (1988) discussed the use of structural equations with latent variables as a modelling tool in travel behaviour analysis. In an application he analysed the causal relationships between mobility, car ownership and income (Golob 1989). Income was measured in terms of three categories. Trip generation was defined as vehicular trips for three modes.

The simultaneous equation model of a joint travel distance and car ownership as a function of income by Golob and van Wissen (1989) for the same data had a novel feature in its specification consistent with the scale of endogenous variables. Income and car ownership levels were measured in terms of ordered categories and treated as ordered probit response variables and modal distances were treated as censored continuous variables subject to a tobit transformation. This research represented an extension of the linear structural equations to non-normal variables. The non-linear transformation of categorical and censored dependent variables was included as a sub-model. The model had only endogenous variables and was estimated using the LISCOMP program (Muthen 1987).

The exogenous variables were included by Golob (1990) in the longitudinal structural model formulated to establish the interrelationships in travel time expenditure by mode and car ownership. The exogenous variables included the household characteristics such as income, household size and age composition, number of workers and the number of drivers. In addition residence location was also included as a static variable, important in explaining travel time.

Following the line of research described above a structural equation model with latent variables will be used in our study of the change in travel demand over time. Our primary objective will be to understand the causal relationships among concepts underlying the theory of travel demand and to estimate these relationships in time and over time by using the observed indicators of these concepts.

## Data to be used in estimation

The 1991 Sydney Travel Study data will provide a current data base for the entire Sydney public transport network (and the highway network). Major household surveys will provide information from 70,000 households throughout the Sydney Metropolitan Area giving details of travel patterns and socioeconomic status. Although the sampling unit for this survey is the household, personal travel details of each individual member of the household have been obtained. Initial transit demand can be estimated using details of the reported journeys and the modes chosen and with the network data providing the required time and cost information about alternative modes.

The travel survey years were chosen to coincide with the census year thus Australian Bureau of Statistics census data can be used to augment our socioeconomic information, as will planning department land-use data. This information will be combined with available data from the Sydney Travel surveys of 1971 and 1981 to allow for time dependence. Whilst this data is not true panel data rather a repeated cross section, since different households were interviewed in each wave, Goodwin et al. (1987) have shown that such data can be formulated in a dynamic way with lags, inertia, and asymmetry.

The large volume of data will allow separate models to be estimated for sub regions as well as for the Sydney Metropolitan Area as a whole. Regions with similar socioeconomic and geographic characteristics will be compared to search for similarities and differences in the levels of demand and seek the factors which cause them.

## Method

### Theoretical justification

An interdependent system of equations describing technological and behavioural relationships among variables in the form of a linear simultaneous equation system has limitations especially in social research. The limitation arises when the structural model is formulated for observed (or measured) variables because errors of measurement may exist in the variables. Conclusions derived from such models can be varied by increasing or decreasing error variance in exogenous variables arbitrarily (Bentler 1983).

The structural model can be formulated in terms of unobserved (latent, or conceptual) variables (or common factors) to identify the "true" causal relationships (Herting 1985). In other words the model describes the "assumed causal structure". Latent endogenous variables are related to other latent endogenous variables and to the latent exogenous variables. The latent variables are assumed to be without measurement errors; however, there will be equation errors which indicate that the latent endogenous variables are not "perfectly predicted by the structural equations" (Long 1983).

In addition to the structural model describing the causal relationships of the process being modelled, a measurement model is needed to relate the latent variables, both exogenous and endogenous, to their observed indicators. These relationships will reflect the measurement errors in the observed variables.

The interrelationships among observed variables as indicated by their covariances are known but they are contaminated by errors in measurement. The aim is to explain these interrelationships in terms of relationships among the observed and latent variables (Long 1983). If a simultaneous solution of the two models can be found then both the underlying causal relationships and observed relationships can be explained.

## Components of a general model

The model described above with two components, structural and measurement, is called a Covariance Structural Model (CSM). It will be our modelling tool in analysis of travel demand in general and transit demand in particular. Recent advances in estimation techniques allow discrete and other non-normal variables to be included in the model. For instance, non-normal endogenous variables such as dichotomous, ordered polychotomous, or censored continuous variables are accommodated by transformation to normal continuous variables in LISREL. In LISCOMP, limited and ordered categorical endogenous variables are handled directly. The ESQ program (Bentler 1985) performs normal, elliptical, and distribution free estimation. The components of a general model are shown in Figure 1 and explained below.

Assume that in the process of change in travel demand behaviour which we attempt to model the "true" nature of the process is described by common factors (or latent variables).

Let N be the matrix of two latent endogenous variables: N(1) and N(2).
Let E be the matrix of three latent exogenous variables: E(1), E(2) and E(3).
Let § be the matrix of the equation errors.

As an example we can assume some causal relationships: say, E(1) has an effect on N(1) and E(2) has effects on N(1) and N(2); E(3) affects only N(3). We also can assume that there is a two-way cause and effect relationship between N(1) and N(2); and that E(1) and E(3) are correlated. This inner part of the above path diagram (Figure 1) is the structural model for the underlying variables which are not contaminated by measurement errors; they are unobserved variables. There are only equation errors which indicate that the endogenous latent variables are not perfectly explained by the exogenous latent variables.

The outer part of the path diagram shows the assumed relationships among the latent variables and their indicators, the observed variables.

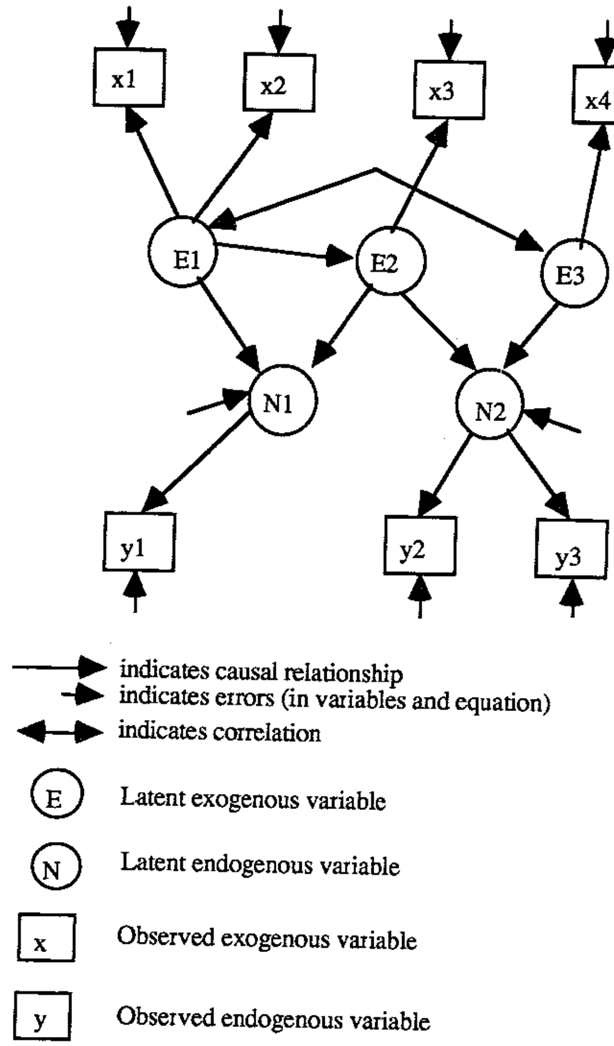


Figure 1 Components of a general causal model with hypothetical relationships

Let **x** be the matrix of four observed exogenous variables: x(1), x(2), x(3), x(4);
Let **y** be the matrix of three observed endogenous variables: y(1), y(2) and y(3);
Let **u** be the matrix of measurement errors in x's;
Let **v** be the matrix of measurement errors in y's.

This is the measurement model with two components: the first indicating the relationships among the latent exogenous variables **E** and the observed exogenous variables **x** ; the second indicating the relationships among latent endogenous variables **N** and the observed endogenous variables **y**.

In this example it is assumed that E(1) is related to two indicators x(1) and x(2); E(2) is related to x(3) and E(3) to x(4). Similarly N(1) is measured by y(1) and N(2) by y(2) and y(3). The small arrows in the diagram attached to the x's and y's are the measurement errors and those attached to N's are equation errors. The longer arrows in the diagram indicate cause-effect relations with their directions. The double arrow indicates correlation between two variables.

This model can be formulated as:

Structural model: $\mathbf{N} = \mathbf{B}\,\mathbf{N} + \mathbf{T}\,\mathbf{E} + \S$

where **B** is the matrix of relationships among the endogenous latent variables;
**T** is the matrix of relationships among exogenous and endogenous latent variables.

Measurement model: $\mathbf{x} = \Delta(\mathrm{x})\,\mathbf{E} + \mathbf{u}$

$$\mathbf{y} = \Delta(\mathrm{y})\,\mathbf{N} + \mathbf{v}$$

where $\Delta(x)$ is the coefficients (also called loadings) matrix, indicating how a change in an exogenous latent variable affects an observed variable;
$\Delta(y)$ is the loadings matrix for the endogenous latent variables and endogenous observed variables.

Given the observed data x's and y's, the variances and covariances can be calculated. Let **S** be the calculated covariances among the observed variables. The covariances among observed variables can be expressed in terms of the parameters of the model:

1. Loadings matrix between x's and E's, $\Delta(x)$;
2. Loading matrix between y's and N's, $\Delta(y)$;
3. Covariance matrix for error term u's represented by **Ø**(x);
4. Covariance matrix for error term v's represented by **Ø**(y);
5. Covariance matrix for the latent exogenous variable E's represented by $\Omega$

Covariances among the latent endogenous variables are expressed in terms of the above parameters. Let the covariances among the observed variables be represented by $\Sigma$.

The aim of the model is to estimate a set of parameters which will minimise the difference between the estimated covariances $\Sigma$(estimated) and the calculated covariances S. At the specification phase of the model constraints will be imposed on the parameters in order to achieve identification. Assuming the model is identified it may be estimated using software such as LISREL, LISCOMP, ESQ, or COSAN (McDonald 1980).

## A dynamic model structure proposal for Sydney transit demand

To study the underlying factors which affect demand for travel at the micro-economic level we propose a structural model of the type discussed above. We suggest that there are three types of factors:

a. Mobility
b. Socio-economic characteristics
c. Individuals' land-use characteristics

These are latent variables. Mobility and land-use are endogenous and the socio-economic characteristics variable is exogenous. Socio-economic characteristics will affect both mobility and land-use. Each of these latent variables may have several observed indicators. Some of the possible sets of indicators are listed below:

*Mobility indicators:*

a. Public transit trips
b. Car trips
c. Public transit travel distance
d. Car travel distance
e. Car ownership
f. Public transit and car attributes

*Socio-economic characteristics:*

g. Income
h. Household size
j. Life cycle of household
k. Number of workers
l. Number of drivers
m. Age
n. Gender
o. Education level
p. Company car use
q. Employment job category
r. Employment status
s. Ethnic background

*Land-use characteristics:*

t. Residence area per individual
u. Employment area per individual
v. Residence location
w Employment location

z. Home business / teleworking

A path diagram which indicates an example of a set of causal relationships and their directions is shown in Figure 2 below. The ellipses in the diagram indicate endogenous (**N**) and exogenous latent variables (**E**). The arrows show the causal relationships and their directions.

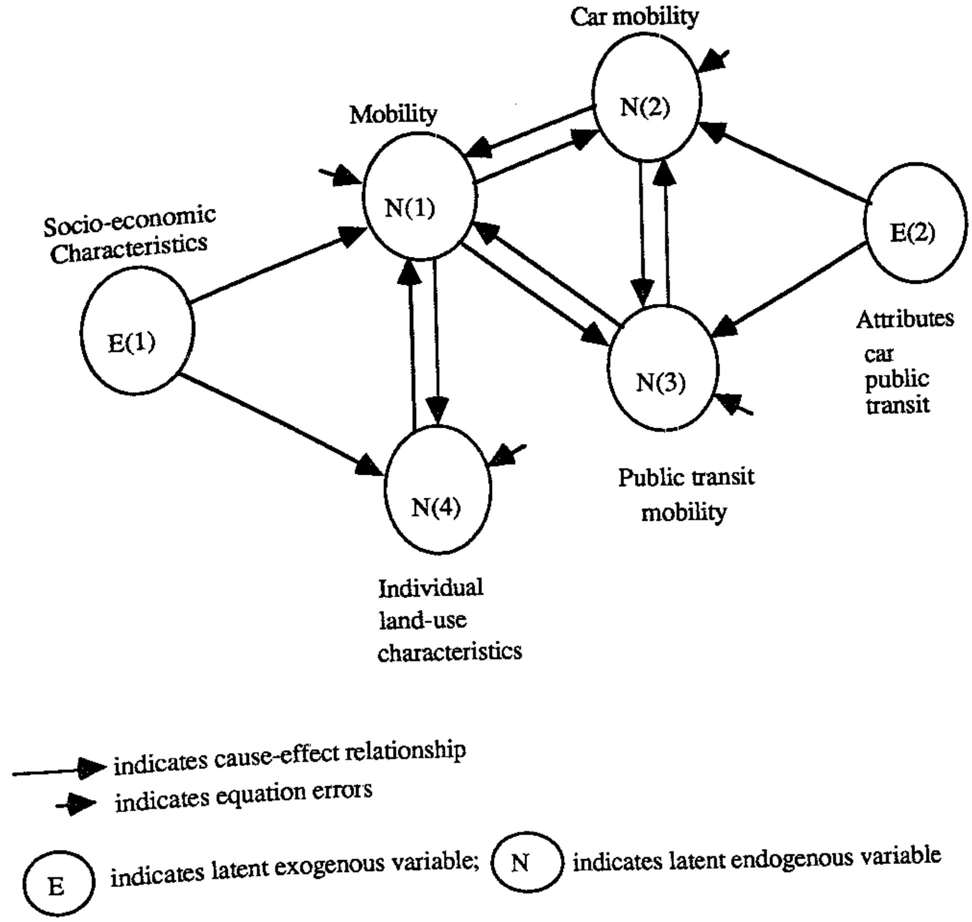


Figure 2 Components of the structural model with assumed relationships

Structural model

As shown in Figure 2 the latent variables and the assumed relationships among them are as follows:

N(1): Mobility, endogenous

N(2): Car mobility, endogenous
N(3): Public transit mobility, endogenous
N(4): Individuals land-use characteristics, endogenous
E(1): Socio-economic characteristics, exogenous
E(2): Attributes, exogenous

E(1) affects N(1) and N(2); E(2) affects N(2) and N(3).
N(1) and N(4) affect each other.
N(1), N(2), and N(3) affect each other

Measurement model

The relationships among the latent variables and the observed variables are explained and indicated in the path diagrams for each sub-model (Figures 3 to 5):

*Mobility*

Possible observed endogenous variables are:

| | | |
|---|---|---|
| y1 | Public transit trips | (number per year) |
| y2 | Car trips | (number per year) |
| y3 | Public transit travel distance | (kms per year) |
| y4 | Car travel distance | (kms per year) |
| y5 | Car ownership | (number of cars) |

Possible observed exogenous variables are:

| | | |
|---|---|---|
| x1 | Public transit accessibility | |
| x2 | Public transit total travel time | (minutes per trip) |
| x3 | Public transit cost | (dollars per trip) |
| x4 | Car travel time | (minutes per trip) |
| x5 | Car travel cost | (dollars per trip) |

The path diagram for the mobility sub-model is shown in Figure 3

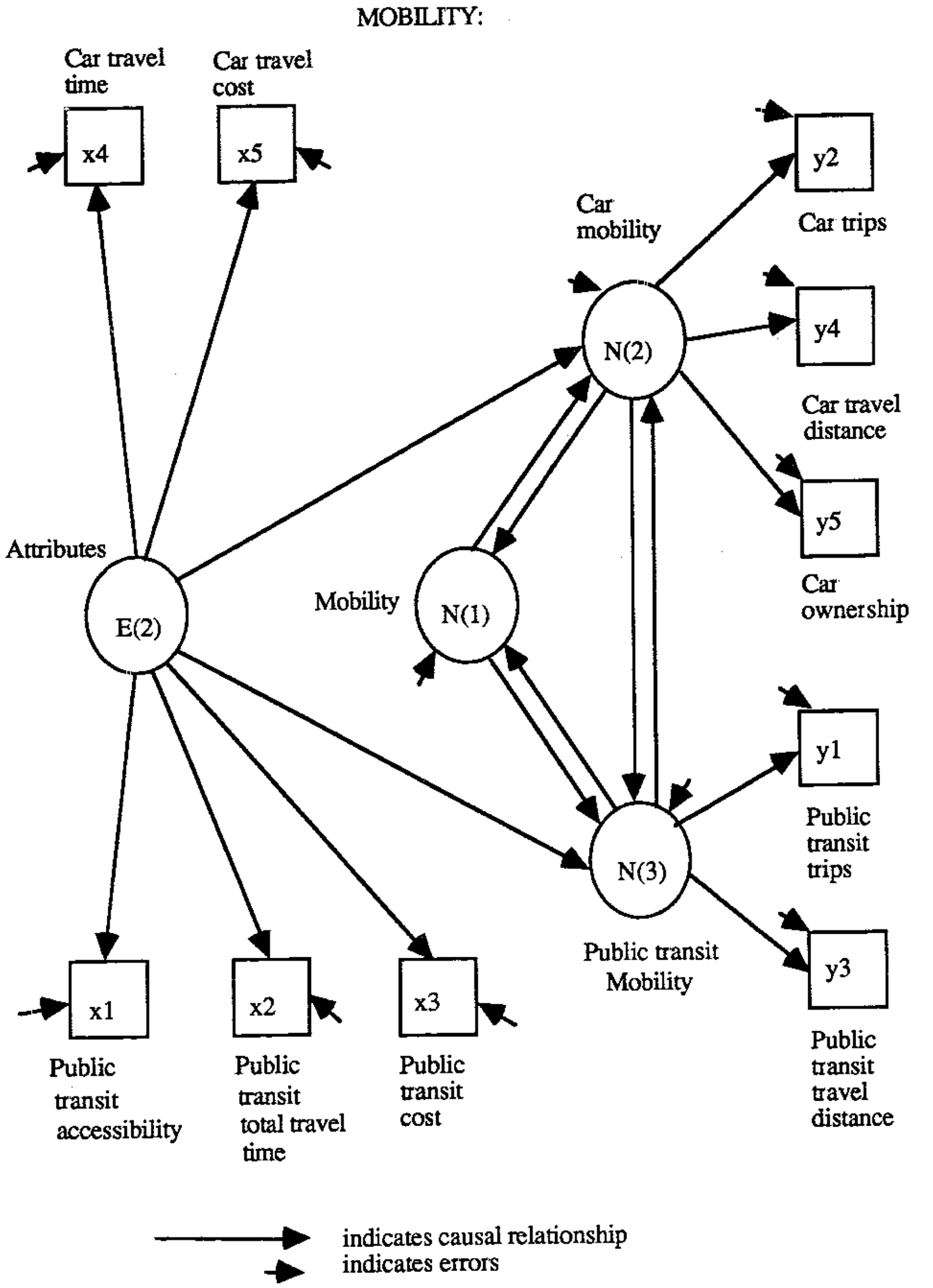


Figure 3 Components of the measurement model: mobility sub-model with assumed observed variables for each latent variable

*Socio-economic characteristics*

| | | |
|---|---|---|
| x6 | Income | (several categories) |
| x7 | Household size | (number) |
| x8 | Life cycle of household | (categories) |
| x9 | Number of workers | (number) |
| x10 | Number of drivers | (number) |
| x11 | Age | (categories) |
| x12 | Gender | |
| x13 | Education level | (categories) |
| x14 | Company car use | |
| x15 | Employment job category | |
| x16 | Employment status | |
| x17 | Ethnic background | |

The assumed relationships of these variables are shown in a path diagram in Figure 4.

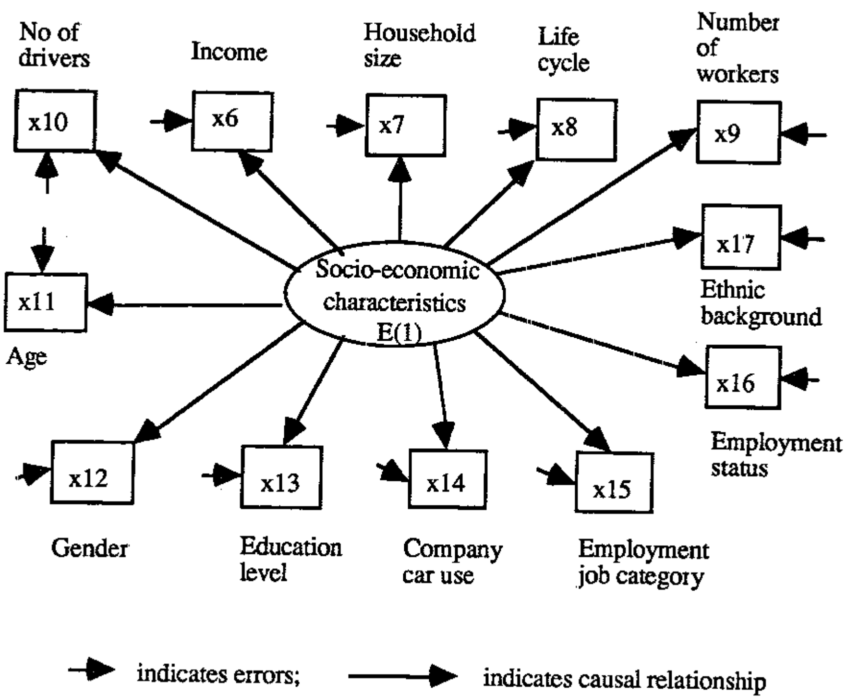


Figure 4 Components of the measurement model: socio-economic characteristics with the assumed observed variables

*Land-use characteristics of individuals*

"Our cities are in transition from industrial to information economies, and the nature and location of work activities are changing" (Brotchie 1991). "The metropolitan transport task is currently about 60% of that of a single-centred city, but the travel patterns have also changed away from just the radial arterials to a more complex multi-centred system less relevant to the existing radial public transport networks" (Brotchie 1992). As stated clearly by Brotchie, land-use has been changing and with the use of data spanning 20 years the effect of land use on transport demand can be examined. Land use is an underlying causal factor in a bi-directional relationship with the other endogenous variable, mobility. Since our model is of individual transport choice, we intend to specialise land-use characteristics, such as medium density housing, to the individual:

| | |
|---|---|
| y6 | Residence area per household |
| y7 | Employment area per individual |
| y8 | Residence location |
| y9 | Employment location |
| y10 | Home business / teleworking |

A path diagram for this sub-measurement model is shown in Figure 5.

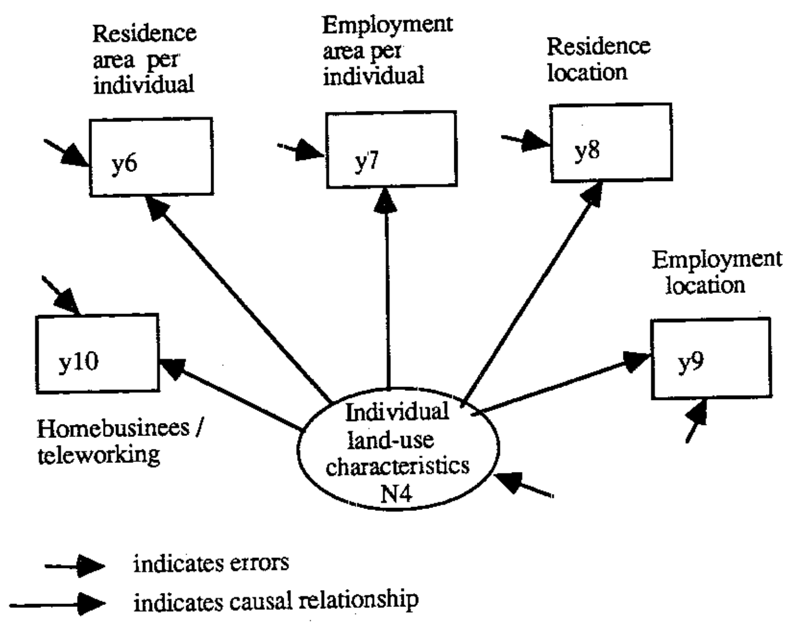


Figure 5 Components of the measurement model: individuals' land use characteristics with assumed observed variables

## Conclusion

This paper documents our current approach to development of a dynamic model of transit demand. In essence we propose a general covariance structural model comprised of structural and measurement submodels. We assume that the "true" nature of changes in travel behaviour is described by latent variables which are hence not subject to measurement error. The structural sub-model is a causal model of latent variables. These latent variables defined as mobility, socio-economic effects, individuals' land-use characteristics are related to observed variables in the measurement sub-model.

Whilst we are not yet certain of our final model structure we are quite certain of both the need for such a model and the inherent complexities it must encompass. Personal experience tells us that choice of transit mode is linked with a combination of circumstances involving residential location and work opportunities together with individual preferences. Moreover these circumstances and opinions are not changed in an instant. Even after decisions for change are made time is required to implement change especially if significant infrastructure is required

Transport infrastructure is expensive to provide. A good model of transit demand can form the basis for a forecasting tool which enables planners to make good estimates of the demand for a proposed public transit service and so avoid expensive mistakes. Better still we should go beyond simply forecasting demand for transport to seeking ways of actually altering that demand by designing services to encourage optimum demand. This should include short run as well as long run strategies. Since our model will be designed for transferability to other major urban centres in Australia, it will allow encouragement of appropriate public transit to suit the urban Australian context.

## Acknowledgement

The work reported in this paper was funded by a research grant from the Australian Research Council.

## References

Bentler, P M (1983) Simultaneous equation systems as moment structure models *Journal of Econometrics* 22, 13-42

Bentler, P M (1985) *Theory and Implementation of ESQ: A Structural Equations Program, Version 2* Los Angeles: BMPD Statistical Software

Brotchie, J (1991) The changing structure of cities *Urban Futures* Special Issue No.5 Canberra: Department of Health, Housing and Community Services, 13-27

Brotchie, J (1992) Developments in urban transport Paper presented at *Planning Co-ordinating & Funding Urban Transport Conference* Sydney: IIR Conferences

Clarke, M, Dix, M and Goodwin, P (1982) Some issues of dynamics in forecasting travel behaviour - a discussion paper *Transportation* 11, 153-172

Golob, J M, Schreurs, L J M and Smit, J G (1985) The design and policy applications of a panel for studying changes in mobility over time in *Behavioural Research over Time,* Utrect (The Netherlands): VNU Science Press, 81-85

Golob, T F (1988) Structural equation modelling of the dynamics of travel choice behaviour Paper presented at *the Oxford Conference on Travel and Transportation*, July, Oxford, England

Golob, T F (1989) The causal influences of income and car ownership on trip generation by mode *Journal of Transport Economics and Policy* XX111, 141-162

Golob, T F (1990) The dynamics of household travel time expenditures and car ownership decisions *Transportation Research A* 24A(6), 343-463

Golob, T F and Meurs, H (1987) A structural model of temporal change in multi-modal travel demand *Transportation Research A* 21A(6), 391-400

Golob, T F and van Wissen, L (1989) A joint household travel distance generation and car ownership model *Transportation Research B* 23B, 471-491

Golob, T F, van Wissen, L and Meurs, H (1986) A dynamic analysis of travel demand *Transportation Research A* 20A(5), 401-414

Goodwin, P B (1989) Family changes and public transport use 1984-1987 A dynamic analysis using panel data *Transportation* 16, 121-154

Goodwin, P B, Dix, M C, and Layzell, A D (1987) The case for heterodoxy in longitudinal analysis *Transportation Research A* 21(4/5), 363-376

Goodwin, P B and Mogridge, M J H (1981) Hypotheses for a fully dynamic model of car ownership *International Journal of Transport Economics* VIII(3), 313-326

Hensher, D A (1988) Model specification for a dynamic discrete-continuous choice automobile demand system in discrete time in Golledge, R and Timmermans, H (eds) *Behavioural Modelling in Geography and Planning* London: Croom Helm, 451-476

Hensher, D A and Smith, N C (1990) Estimating automobile utilisation with panel data: an investigation of alternative assumptions for the initial conditions and error covariances *Transportation Research A* 24A, 417-426

Hensher, D A, Smith, N C, Milthorpe, F W and Barnard, P O (1992) *Dimensions of Automobile Demand* Amsterdam: North-Holland

Hensher, D A and Wrigley, N (1986) Statistical modelling of discrete choices in discrete time with panel data in *Behavioural Research for Transit Policy* Utrect (The Netherlands): VNU Science Press, 97-116

Herting, J R (1985) Multiple indicators in LISREL in Blalock, H M (ed) *Causal Models in Social Sciences* New York: Aldine Publishing Company, 263-320

Joreskog, K G (1978) Structural analysis of covariance and correlation matrices *Psychometrika* 43, 443-477

Joreskog, K G and Sorbom, D (1984) *LISREL VI: An Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Square Methods* Mooresville, Indiana (USA): Scientific Software Company

Long, J S (1983) *Covariance Structure Models - An Introduction to LISREL* Sage University Paper Series No. 34 Beverly Hills and London: Sage Publications

McDonald, R P (1980) A simple comprehensive model for the analysis of covariance structures: some remarks on applications *British Journal of Mathematical and Statistical Psychology* 33, 161-183

Muthen, B (1984) A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators *Psychometrika* 49(1), 115-132

Muthen, B (1987) *LISCOMP: Analysis of Linear Structural Equations with a Comprehensive Measurement Model* Mooresville, Indiana (USA): Scientific Software Company

Newman, P W G, and Kenworthy, J R (1989) *Cities and Automobile Dependence* Brookfield: Gower Technical