A STRUCTURED-LOCUTI MODEL OF AUTOMOBILE ACQUISITION AND TYPE CHOICE

DAVID A. HENSHER
Associate Professor of Economics,
Macquarie University.

TRACEY MANFIELD
Research Associate,
School of Economic and Financial Studies,
Macquarie University.

ABSTRACT: A dominant theme in the literature on automobile ownership has been the estimation of aggregate motor-vehicle demand. Recognition of the structural and estimation deficiencies of the aggregate models (predominantly stock-adjustment models) as well as the need for policy-sensitive vehicle-type choice predictions, particularly as a consequence of rapid and regular fuel price increases (in money terms), has led to the development of individual discrete-choice models of household vehicle ownership.

This paper outlines the first empirical study to model the joint decision of auto-acquisition choices and auto-type choices. It faces up to the decision process that results in a vehicle being in the household. A major emphasis is the demonstration of the usefulness of individual-choice models in the sensitive arena of energy policy and vehicle design.
AUTOMOBILE ACQUISITION AND TYPE CHOICE

INTRODUCTION

This paper develops a household-level set of discrete-choice models to examine the auto acquisition and type choices of a sample of Sydney households. Such a model framework is arguably appropriate to provide advice on the likely consequences of energy-related pricing and non-pricing policies; especially the influence of petrol price increases (both absolute and expectational) and auto fuel efficiency on the composition of the household vehicle fleet. The data is drawn from a pilot survey undertaken in 1980. The pilot data set forms the basis for the development of a panel survey instrument, to be administered annually on a fixed set of households. The pilot data is a single cross-section but used in a dynamic sense - we model change in stock, not absolute stock.

The paper is organised as follows: a brief review of the main emphases of extant auto models is given; then the structured-logit model is presented with special emphasis given to the parametric requirements for consistency with utility maximisation. The next section outlines the models, then we present the analyses and discuss the findings. The concluding section outlines directions for continuing research.

THE EXTANT LITERATURE ON AUTO-DEMAND MODELLING: A SNAPSHOT VIEW

There is no shortage of literature review on modelling the set of automobile demand/choice decisions (e.g. Mellman 1979, Ayres et al. 1976, Train 1979, Tardiff 1980, Charles River Associates 1980, Train 1980, Manski 1981a), and extensive critiques of particular models (e.g. Charles River Associates 1980 on the Wharton Econometric Forecasting Model, and Train 1979 on the Chase Econometric Associates Model). We limit discussion to a snapshot view of

1. Acknowledgements: We wish to thank Vicki LePlastrier, Dick Filmer, Royce Ginn, Bob Mellman, Tim Tardiff, Ken Train, Joel Horowitz, Hayden Boyd, Bill Tye, Len Sherman, Sean Manefield, Dick Westin, Charles Manski and Terry Johnson for discussions on auto-choice modelling.

Earlier versions were presented to Seminars at the University of New South Wales and Monash University.
the two dominant approaches - the aggregate (predominantly stock adjustment) models and the discrete-choice individual household models. Train (1980) discusses other approaches (especially methods based on stated preferences and non-compensatory choice situations). The comments below relate to existing models. It is clear that both approaches could be improved and have a role in more diversified contexts.

The Aggregate Approach

Aggregate auto-demand models are predominantly single equation short-run time series models designed to explain annual variations in auto sales (e.g. Juster and Wachtel 1972, Suits 1958, Wharton Econometric Forecasting Associates 1977). These models assume a desired stock of autos \( A^* \) which is a function of economic variables such as income, price of autos, perceptions of income risk. Households make additions to stocks to reduce some constant proportion of the gap between the desired and actual auto stocks each period:

\[
A_t - A_{t-1} = \alpha(A^* - A_{t-1}) \quad 0 < \alpha < 1
\]  
(1)

Replacement acquisitions are assumed equal to some constant proportion of their stock each period:

\[
R_t = \delta A_{t-1}
\]  
(2)

and gross purchases of autos are equal to net acquisitions plus replacement acquisitions:

\[
X_t = (A_t - A_{t-1}) + R_t = \alpha(A^* - A_{t-1}) + \delta A_{t-1}
\]  
(3)

The majority of the applications of this model either assume proportionality between flow of services and stock or the desired stock is based on economic variables with a demand function for the auto stock. This model form was originally developed for use in large macroeconomic models such as that of Houthakker and Taylor (1966), where little attention was given to each commodity, and one requirement was simplicity, as found in a common (commodity) model. Experiences with this model form are varied. A recent review of USA studies concluded that

"Simple and straightforward models proved to be relatively accurate predictors of size of the aggregate market particularly when markets changed slowly" and

"the purpose of these models was primarily limited to explaining changes in aggregate vehicle sales and industry revenue due to national demographic growth, economic growth and business cycle" (Charles River Associates, 1980)
In contrast, other writers have reservations as to the value of the simple (essentially stock adjustment) models. Armstrong and Odlung-Sme (1978, 1979) and Haig and Gregory (1970) argue that such models fail to specify depreciation unambiguously and demand does not appear to be very sensitive to changes in exogenous variables.

Armstrong and Odlung-Sme proposed and applied a non-stock adjustment model, emphasizing the replacement decision; the main influences on the demand for new cars being the need to replace ageing cars. They distinguished normal replacement demand and abnormal replacement demand (or changes in the timing of normal purchases). The former was determined by the weighted sum of part purchases and the level of exogenous variables (e.g. income, price, liquidity, hire-purchase regulations), the latter was determined by changes in these exogenous variables.

Many recent studies, however, have preserved the stock-adjustment framework, improving its realism by removing its simplicity. For example, Grieve (1980) and Filmer and Mannion (1980) have refined the model to enable both auto services and units of the durable good to be objects of choice. This enables the model to include a variable rate of service flow from a given auto stock. Wykoff (1973) and subsequently Johnson (1978) argued that different vintages are not weighted perfect substitutes and hence are not aggregatable commodities. They proposed the superior goods model in which new car services represent the demand for a unique commodity, measured independently of existing stock of used cars. Wykoff's results show that the stock adjustment approach incurs aggregation bias in estimation of demand function parameters. This superior-goods hypothesis has recently been applied by Greene (1979). There are, however, other ways of handling 'newness' (especially the individual discrete-choice approach - see below), supported by Wykoff's statement (1973, 388) that newness is predominantly a function of attributes of utility (e.g. feeling of freshness, pride of driving newest models, avoiding dealing with used car salesmen). Such dimensions can be (and some have been - Ohta and Grilliches 1975) incorporated in hedonic quality indices, readily obtainable from the auto-type discrete-choice model presented below (see Truong 1981, Hensher and Truong 1981). They enable generalization of vintage beyond the binary assumption of superiority.

Westin (1975) proposed a further modification, on the premise that individual purchasers of consumer durables enter the market at infrequent intervals. Thus the assumption that current replacement demand equals current depreciation on the existing stock can no longer be justified. Loss of service on present stock is not translated into effective demand for new stock. The Westin model, a discretionary replacement model, reinterprets the role of substitution in use between old and new durables and explicitly distinguishes between planned and discretionary replacement behaviour. The model is
The aggregate models assume a representative consumer with deterministic behaviour, with a demand function assuming in its strictest sense a consumer selecting amongst continuous levels of quantity of homogeneous goods. This framework is not well suited to a consumer durable which displays discrete decisions, despite the attempts to modify some of the strictness, especially homogeneity.

The assumption of a perfect used auto market, with a discounted annual depreciation user cost as the appropriate measure of price; or a purchase price as relevant in both used and new auto markets.

Early stock adjustment models assumed the automobile as a homogenous good, whereas the 1970's versions (e.g. Blomquist and Haessel 1978) permitted some heterogeneity by defining autos by size, weight and fuel efficiency, and dropping the assumption of perfect substitutability between new and used autos. Although not wishing to support one modelling approach over another when most of the criticism can be accommodated with better data and less aggregation, it is useful to list some of the criticisms of the available aggregate models:

1. The aggregate models assume a representative consumer with deterministic behaviour, with a demand function assuming in its strictest sense a consumer selecting amongst continuous levels of quantity of homogeneous goods. This framework is not well suited to a consumer durable which displays discrete decisions, despite the attempts to modify some of the strictness, especially homogeneity.

2. The assumption of a perfect used auto market, with a discounted annual depreciation user cost as the appropriate measure of price; or a purchase price as relevant in both used and new auto markets.
3. The treatment of scrappage in a mechanistic way, although more recent studies treat it as an economic decision (e.g., Filmer and Talbot 1975, Parks 1977, Greene 1979). A popular way of incorporating scrappage is to hypothesise that the probability of scrappage (or vehicles surviving to age $a$ that fail to survive to $a+1$) is the likelihood of a repair bill being greater than the difference between the value of a working a year old vehicle and its scrap value. This hypothesis, formulated empirically as a logistic scrappage function for make-vintage-age specific vehicles (Parks 1977, Greene 1979) implies that the household is already achieving optimum utility from the flow of services, a questionable assumption. One way of resolving this is to adopt the structured (discrete) choice approach in the next section and link the scrappage function (choice of disposal - repair and sell, scrap, repair and keep, keep) to other relevant auto-choice decisions.

4. The unsatisfactory incorporation of automobile classes to capture product differentiation (Manski 1981a). Each class contains a very heterogeneous set of vehicles. A finer classification is possible, given the data, but at the expense of considerable complexity (Ben Akiva, Manski and Sherman 1980). Where variables such as sales weighted average of vehicle prices is used as in the Wharton model, with heterogeneous autos in a 'class', it can lead to spurious results, since sales are endogenous variables:

"Assume that for exogenous reasons sales of the lower priced mid-size cars increases and sales of the higher priced cars decreases by a like amount. The result will be a fall in the mid-size price index but no change in total mid-size sales at all." (Ben-Akiva, Manski and Sherman, 1980, 5)

5. The relationship between vehicle stocks and auto use decisions has not been adequately addressed.

2. The explanatory variables in Parks (1977) model of the probability that a vehicle will be scrapped are the vehicle's age, the relative cost of repairs, and dummy variables that capture the built-in durability characteristics. The durability of a vehicle is defined in terms of features that affect the pattern of breakdowns and the nature of repairs that occur over the vehicle's lifetime and that ultimately affect the vehicle's lifetime. Parks assumes that the durability parameter is a function of the cost level of new vehicles, the interest rate, and expected level of prices for used cars, for repairs and for scrap.
It is doubtful that aggregate data can be used reliably to separate influences of different vehicle, household and macroeconomic characteristics on the structure of vehicle demand. Cross-section data (with usually state or county as unit of analysis) suffer from confoundment of explanation associated with variance due to definition of aggregate analysis unit and the real sources of variance in behaviour. Time series approaches suffer from the highly aggregate nature of each period's variables, where a variable such as ratio of employment to vacancies has an equal impact on all individuals in a given time period.

Another stream of aggregate models involves treating all consumer demand in a system. Durable goods (which require utility functions for more than one period) can now be incorporated into such systems (McAleeer, Powell, Dixon and Lawson 1979), using the CONDELES model (Consumer Durables Extended Linear Expenditure System). The services flowing from durables are assumed to be proportional to the stocks held, measured in units which recognise both the quality and age aspect of the vintage composition of the stocks held. The main advantage of this approach over the other aggregate approaches is the interrelationships between all items of consumer demand, its disadvantage its complexity.

An alternative approach, individual discrete-choice models, has been proposed in the last few years, partly in response to the growing dissatisfaction with the aggregate models (even in their most sophisticated form), and partly as a logical extension of their increasingly diversified set of applications.

The Individual Discrete-Choice Approach

This approach, couched in a theory of individual probabilistic-rational behaviour, considers a population of decision-makers each selecting a preferred alternative among a finite set of alternatives (Hanski and McFadden 1981, Hensher and Johnson 1981). The theory and model form is discussed in the next section. The approach has recently been applied to auto ownership (e.g. Lerman and Ben-Akiva 1976, Train 1979) and auto type choice (Lave and Train 1979; Hanski and Sherman 1980). These studies used disaggregate data (i.e. individual households), and the multinomial logit model. Lave and Train investigated the probability of selecting an auto type given that an acquisition had occurred, and that the acquisition was a new car. Hanski and Sherman identified the probability of selecting an auto type given fleet size, and given that a review took place and composition change was considered. Hence both studies model type choice conditional on one of the outcomes of a 'mysterious'
Automobile Acquisition and Type Choice

Decision process. Manski and Sherman (1980a) also included auto ownership and disposal decisions; treating household's choice among vehicle types conditional on ownership level. The Manski and Sherman (1980a) model is the most sophisticated to date. In each submodel, the household is viewed as evaluating its current auto holdings annually and updating them as desired. The relative utility of any auto (or auto combination) is assumed a function of such auto attributes as seating capacity, luggage capacity, weight, acceleration, price, fuel efficiency. A search-transaction cost is associated with entering the market. In addition decision-maker characteristics such as household size, age, education, income, number of workers, and residential location condition the utility function.

The abovementioned studies define the decision-making unit (i.e., the individual or household) as the unit of analysis; in contrast, a recent study by Charles River Associates (Boyd and Melvin 1980, Cardelli and Dunbar 1980) uses the logit form of the discrete choice model but on aggregate sales share data. Their model, referred to as the hedonics demand model, allows the parameters to be random, hence aggregating individual demands in accordance with a distribution of tastes. Consequently, the estimated equations for aggregate demand are consistent with a realistic model of individual consumer behavior. In common with the neoclassical approach, the hedonics model has aggregation bias due to defining the unit of analysis as the percentage of sales of model type.

A major appeal of the individual discrete-choice approach is the ability to establish the choice process through imaginative use of data on both alternatives chosen and not chosen, and to structure a complex set of interdependent decisions in such a way that estimation is manageable and meaningful (see next section). Furthermore, the use of individual-specific data greatly enhances the opportunity to understand the real sources of variance in behavior, avoiding the fallacy of composition so common with aggregate data. If the relationships between individual decisions and aggregate demand is understood, then 'extensive' data on individual choice can be used to refine estimates of the aggregate demand function.

Despite these strong positive points, there are many concerns and unanswered problems. In particular, all applications to date use a single cross-section of data, preventing any strong evidence on the stability over time of parameters and range of variables. Indeed, in an area of such radical change (as auto type), we can be greatly concerned about this. Also, we cannot ignore the
supply side, especially the restrictions imposed on imports of currently popular Japanese cars. This has created a situation of disequilibrium which must be reflected in the auto type choice decision (either in deferral of acquisition of a type or in selection of another type).

The simplest econometric form, multinomial logit suffers from the independence-from-irrelevant alternatives property which almost certainly is not satisfied in the auto type choice. However, as discussed below, there are ways of minimizing violation of this property, even though we could not guarantee complete non-violation.

Unresolved issues include (a) the ability of the approach to handle in a tractable manner panel data on a number of inter-related decisions, especially auto acquisition, auto disposal, auto type, auto usage, auto age, auto expenditure, residential location, tenure type, and commuter mode choice and (b) the predictive power of discrete-choice models. It is clear that the individual discrete-choice approach has great appeal. However, whether the aggregate framework could equally accommodate the new set of concerns of government and industry is itself of great importance. These concerns include:

(a) the effectiveness of a policy designed to encourage household auto downsizing (and increased fuel efficient autos).
(b) predictions of response to changes in auto design.
(c) the influence of petrol price increases (level and frequency) on auto stock adjustment and travel patterns (auto, non-auto).
(d) the influence of combining threats of an impending petrol shortage on expectation of future petrol prices and auto type acquisition/disposal decisions.
(e) the implications of the predicted adjustments in the size and composition of the auto fleet on levels of traffic congestion, and air pollution.
(f) the relationship between residential location and auto ownership, hypothesized to be affected by the relative price changes of home ownership and auto ownership (including operating and maintenance costs).
THE STRUCTURED-LOGIT APPROACH

The approach adopted requires a theory of individual choice behaviour, a model to combine the behaviour rules of each individual to explain population choice behaviour whilst maintaining the decision-making unit as the unit of analysis, and an analytically tractable estimable econometric model obtained by the imposition of testable assumptions on the theoretical choice model. This is the basis of the individual discrete-choice modelling approach, now used extensively in the modelling of more traditional travel decisions (e.g. mode and destination choice), and recently applied to the study of auto-related decisions (especially auto-type choice). The literature on discrete-choice modelling is extensive; with comprehensive statements available in Hensher and Johnson (1981), Manski and McFadden (1981) and Manski (1981). We briefly summarise the particular model form, structured logit, used in the current application.

Define a random utility function depicting probabilistic rational behaviour, through the assumption that the underlying preferences of individuals undergo random changes (3). There is widespread evidence (e.g. Koo, 1963) that inconsistencies in behavioural patterns are more apparent than one would like to believe: "...If economic agents reveal at all a certain consistency of behaviour, this consistency is at best of a probabilistic nature" (Hildenbrandt, 1971, 414).

\[ u(x) \equiv u_{q}(x) \text{ with a probability } P_{q}, \quad q = 1, \ldots, Q \quad (4) \]

where \( u_{q}(x) \) is a particular form of the utility expression, an individual behaviour rule, \( x \) is a vector of attributes, \( u(x) \) is a general (or model) form of the utility expression, a set of individual behaviour rules, and \( Q \) is the number of observations. The form (4) is too general; hence additional assumptions are imposed; in particular the form of the function is assumed unchanged but the momentary value is subjected to random shocks:

\[ u(x) = v(x) + \varepsilon(x) \quad (5) \]

where \( v(x) \) represents the deterministic element and \( \varepsilon(x) \) the random element of the function. A family of random utility forms have been developed, each differing according to the distributional assumption imposed on \( \varepsilon(x) \).

In contrast to the decision process undergoing random changes, in permitting the latter source of randomness, we introduce the complex issue of stochastic transitivity.
In accordance with the maintained assumption of utility maximisation, an individual will select alternative $i$ if the utility to be gained from consumption of services originating from $i$ is greater than the utility to be derived from all other alternatives in the set; i.e. (dropping $q$ subscript)

$$P_i = P[u(x_i) > u(x_j), \forall \ j \in J, i \neq j]$$

(6)

$$P[u_i(x_i) - u_j(x_j) > e_j(x_j) - c_j(x_j), \forall \ j \in J, i \neq j]$$

(7)

The distributional assumptions imposed on $e_j(x)$ are:

- $e_j$ are independently and identically distributed
- $e_j$ are distributed as double exponential
- $E(e_j) = 0$, and $Var(e_j) = \sigma^2$

It follows from these assumptions that $c_i - e_i$ are distributed logistic (Dowd and McFadden 1975, 63), $E(u_i - u_j) = v_i - v_j$, and

$$Var(u_i - u_j) = \frac{\sigma^2}{\lambda^2} = \frac{\sigma^2}{2}$$

The resultant closed-form model is known as the basic multinomial logit (BNLL) model:

$$P_i = \frac{\exp[\lambda v_i]}{\sum_{j=1}^{J} \exp[\lambda v_j]}$$

(8)

$\lambda$ as a scale parameter is the only parameter characterising the variances of differences in utility. This means that in the BNLL model the variance between any two alternatives is independent of which alternatives are being compared. This implies the well-known 'independence from irrelevant alternatives' (IIA) property, which as immortalised in the red bus-blue bus fable is both a potential strength (because of simplicity) and potential weakness (because of lack of realism) of the BNLL model. The discussion of the IIA property is well documented in many sources (e.g., Hensher and Johnson 1981; McFadden, Train, and Tye 1977); suffice to say that there are procedures available to test for its violation and remedies available when violation is detected. (6) One approach designed to minimise violation of the IIA property while still maintaining the analytical and estimation niceties of the logit framework is the 'structured logit' (SL) model. It takes advantage of the existence of a single dispersion parameter, and defines a structural relationship between decisions (or elements of a 'single' decision) such that $\lambda$ remains constant across alternatives in a given choice set, while allowing it to vary across choice sets. The idea is not new, but has only recently been formalised in the context of a theory of behaviour such that the elements of the SL model are consistent with utility maximisation. We now turn to a discussion of the SL model, since we adopt it in the modelling of auto acquisition choice and auto type choice, where we assume that $\lambda$ associated with each of the two decisions is invariant within each decision but varying between decisions.

4. Only two studies can be cited that actually tested for violation, out of the hundreds of applications of the BMNL model.
AUTOMOBILE ACQUISITION AND TYPE CHOICE

The essence of the SL model is that it enables us to decompose a complex set of decisions (which may be jointly decided) into a sequential recursive set of decisions. Williams (1977), Daly and Zachary (1978), Ben-Akiva and Lerman (1979) and McFadden (1979) have proven that \( \lambda \) is directly linked to the parameter of the logarithmic sum of the denominator of equation (8), and is interpreted as an index of the similarity of alternatives provided it is bounded by 0 and 1, a necessary and sufficient condition for the model to be consistent with utility maximisation. This enables use of the dispersion parameter in testing for sequence and structure.

The SL model, for two decisions (or a two-level decision) is

\[
\begin{align*}
p_{at} &= \frac{\exp[\lambda_t (v_t + v_{ta})]}{\sum_{t' \in T} \exp[\lambda_t (v_{t'} + v_{t'a})]} \cdot \frac{\exp[\lambda_a (v_{at} + \ln T \sum_{t' \in T} \exp[\lambda_t (v_{t'} + v_{t'a})])] a \in A}{\sum_{a' \in A} \exp[\lambda_a (v_{a'at} + \ln T \sum_{t' \in T} \exp[\lambda_t (v_{t'} + v_{t'a'})])] a' \in A} \\
&= p_{t|a} p_a
\end{align*}
\]

where \( t \) and \( a \) refer to the auto-type and acquisition decision alternatives:

\[
\lambda_t = \frac{r}{\sqrt{2} \sigma_A}
\]

\[
\lambda_a = \frac{r}{\sqrt{2} \sigma_A} = \frac{r}{\sqrt{2} (\sigma_A^2 + \sigma_t^2)} \frac{1}{2}
\]

(Williams 1977, 316)

For equation (9) to be consistent with utility maximisation; then

\[
\lambda_a < \lambda_t \tag{10}
\]

\[
0 \leq \lambda_a \leq \lambda_t \leq 1 \tag{11}
\]

Intuition tells us that for alternatives with relatively more similarity, \( \lambda \) is greater than for alternatives with less similarity. This is because as \( \text{var} \left( \varepsilon_i - \varepsilon_j \right) \) increases, \( \lambda \) decreases and vice versa. Why should \( \lambda_a \) be less than \( \lambda_t \)? The variances of decisions made earlier in the hierarchy must be greater than variances of decisions made later. This follows from the model framework since we assume that individuals take account of the 'expected maximum utility' (EMU) at earlier levels and these in turn suggest a wider range of possible dispersion between decisions taken at these levels.
In the only element of a dispersion parameter is $\lambda_a / \lambda_t$ the parameter 
attached to the second term in the numerator and denominator, referred 
to as the expected maximum utility of the auto-type choice process.

Proof of this is given in Ben-Akiva and Lerman (1979). In logit estimation, 
the dispersion parameter in the higher order decision is usually 
normalised; hence $\lambda_a = 1$; giving $1/\lambda_t$ as the coefficient of EMU.

The reasoning behind condition (11) is best illustrated by 
interpreting the various magnitudes of the ratio $\lambda_a / \lambda_t$ if there 
is an improvement in the utility of acquisition 1 over a associated 
with auto type $t$, then if

$$\lambda_a / \lambda_t < 0,$$

the overall response will be an increase in the number of 
decisionmakers acquiring a of auto type $t'$

$$\lambda_a > 1,$$

there will be a greater change from $t'a'$ to $t'a$ than from 
t'a to ta. That is, the elasticities will be of the wrong 
sign.

$$\lambda_a / \lambda_t = 0$$

The increase in ta will match precisely the decrease in ta'.

In the first two cases, the effect of the improvement in $\lambda_a$ is an increase 
in selection of $t'$ in a, rather than $t$. These results are unacceptable, 
given the postulate of utility maximisation. When $\lambda_a / \lambda_t = 0$, the SL 
model reverts to a model of two independent BMNL models.

$$0 < \lambda_a / \lambda_t < 1,$$

means there will be a greater change from $t'a'$ to $ta$, than 
the diversion from $t'a'$ to $t'a$.

$$\lambda_a / \lambda_t = 1,$$

the diversion from $t'a'$ to $t'a$ will equal the movement from 
t'a to ta. In this case, the rates of substitution between the 
attributes of different alternatives will be the same, and hence the 
SL model is the same as a simultaneous model.
Equation (9) is the specification of the model to be estimated, with the auto-type choice model estimated first, and the expected maximum utility associated with this choice process being included as an independent variable in the auto-acquisition choice model. While acknowledging that there are other related decisions such as auto disposal, residential relocation and vehicle usage, these are treated as exogenous in the current analysis; and made endogenous in research currently in progress.

The data collected as part of a pilot survey for a larger study, included 400 household interviews in the Sydney Metropolitan Area on a sample, stratified by household fleet size and geographical location. The questionnaire sought extensive details on each automobile regularly parked at the household's address at February 1980, information on changes during the previous 12 months to the number and types of vehicles, the alternative vehicles considered during the period of assessing fleet adjustment, travel by each vehicle during the 12 month period (distinguishing travel in the Sydney area, travel outside of the Sydney area), expectations of petrol price increases up to 1983, and the usual socio-demographic household data. The particular data items obtained are listed in Table 1. Supplementary technical data on each automobile was obtained from the National Roads and Motorists Association (NRMA) of New South Wales.

151 observations were suitable for modelling the two choices selected. Given the pilot nature of the survey, no attempt was made to obtain missing information. Sixty households had acquired one vehicle, and ninety-one households had acquired zero cars during the 12-month period. There is no attempt at representation; this is an exploratory phase in a larger study. The empirical results are of interest; and serve to at least highlight the potential of the econometric approach used.

The auto-type choice model requires definition of alternatives. Like previous studies (e.g. Manski and Sherman 1980), households are assumed to re-evaluate their fleet composition at the beginning of each year, and ensuing decisions remain until the subsequent year. This simplifies the modelling considerably, especially where the output is to be used for prediction. Whether this is an appropriate assumption is unknown; it is possible that in a dynamic-choice context the timing and frequency of such reappraisal periods has an important effect on the outcome (given the influence of period and supply effects). For a single cross-section study, these concerns are added to the other missing dynamic factors. One of the unresolved issues in auto-type choice modelling is the definition of the choice set. Some studies (e.g. Lave and Train 1979) define a fixed set of alternatives (ten size/price categories) and assume all are relevant alternatives for each household. Average levels of attributes are associated with each alternative (e.g. the average horsepower...
Table 1  Data Obtained from Survey

<table>
<thead>
<tr>
<th>Vehicle Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Make, model, engine size, body type, year of manufacture, year household acquired vehicle, financing of acquisition, age at acquisition (new, used), manual or automatic, air conditioning, tow bar, purchase price, maintenance by household or other, expenses (registration, insurance, maintenance, repairs, fuel, tax deductions, fuel source, annual kilometres, type of garaging, insured value, expected maximum price at current disposal, type of insurance, time to dispose, role of vehicle in the acquisition-disposal decision; allocation of Sydney area kilometres between weekday commuting, weekday use in work, non-work weekend use, work weekend use; allocation of Sydney area kilometres between local/adjacent suburb use and to/from central area; kilometre distribution (times/week) in excess of 80 km per day; longest trip within Sydney area (kms, purpose, household participants), details on up to three holidays trips outside Sydney Area (kms, mode, duration away, household participation, touring).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-Vehicle Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel price expectations (1981, 1982, 1983), household composition (number by age, driver's licence, employment - full-time, part-time, students, retired seeking work, full-time housewives), car availability for commuting, workplace fixity, access to each vehicle, alternative mode usage for commuting, household gross income, purchase of petrol by each household member, mode used for shopping journey, plans to relocate residence.</td>
</tr>
</tbody>
</table>
of a subcompact) Other studies (e.g., Manki and Sherman 1980) define the choice set as the chosen auto type and a randomly selected set of twenty-five alternatives from the 600 types available. Unlike the Lave-Train approach, the Manki-Sherman approach maintained each vehicle as unique (i.e., auto-specific attribute levels); however, they had difficulty in identifying which vehicle the household held. To overcome this they defined vehicle class (by make, model, vintage if domestic, and make, vintage if imported), and assumed that the specifications of a given class were those of the top selling vehicle in the class. Both of these studies did not obtain the non-chosen alternatives from the sample of households; hence it was a relatively straightforward exercise to define type choice.

There are two schools of thought, however, in establishing the choice set for each household. One is the approach adopted by the studies referred to above which obtain the non-chosen alternatives independent of the household's perceived set. The other approach involves asking the households to indicate the alternatives they actually considered in the choice process. Both approaches have their strengths and weaknesses. The former approach has a forecasting advantage in that all makes and models are included a sufficient number of times to enable meaningful (statistically) statements on the likelihood of selecting each auto type. However, if this approach fails to establish through the set of independent variables the real impact of constraints on selection amongst the full range of alternatives, it is argued that the direct questioning of individuals is more suitable. If there is any evidence of randomness in the allocations by households, then both approaches should yield similar (if not identical) results. Preliminary assessment using the current data set and a data set using the retrospective longitudinal survey technique (Lecestrier 1981) indicate that if randomness exists, it exists only within a price/size class of auto-types. For example, the alternatives to 'large-imported expensive autos' are 'large-imported expensive autos'.

In the current application we asked households for the alternatives; the result was that the Lave-Train and Manki-Sherman type categories were not appropriate for reasons already mentioned. Within a class of types (e.g., small autos) the correlation between alternatives (through the attribute set) prevented use of existing type classifications. As a consequence (which we argue is likely to be the real situation), two approaches remain; either extensive searching using classification techniques is required to identify new categories, or we test a number of 'simple' dimensions, which have meaning in the context of study aims. The resolution was to adopt a fuel-related criterion. We used exploratory data analysis as the basis of establishing the type categories, such that the correlation between the attributes on each alternative used in hypothesis testing is minimised; and that the categories...
were meaningful in a policy setting. Three categories enabled definition of operational choice sets, which contained a random allocation of price/size vehicles across categories. These are:

1. Low fuel consumption autos (up to 10.49 litres per 100 kms)
2. Medium fuel consumption autos (10.50 to 13.99 litres per 100 kms)
3. High fuel consumption autos (14 litres per 100 kms and over).

Each household in the estimation sample had a choice set containing two or three alternatives.

As a final point on type choice, we have assumed a single level in the decision hierarchy; however, it may be more appropriate to consider a nested structure within the type-choice decision; modelling within a class (e.g., small) separately and then modelling between classes; using the EMU variable to link the levels (see Fig. 1). The only concern would be the failure of the first stage estimation to provide meaningful (if any) results, given the likely high correlation between alternatives. A factor analysis stage is probably required instead of stage 1.

Figure 1: An Alternative Type Choice Structure
In contrast to the type-choice model, the acquisition-choice model is straightforward in the current application. We have limited the model to situations where a household acquires only one vehicle during the twenty-month period. This restriction was imposed by the data. Hence the complete model structure considers the probability of a household acquiring one auto or not acquiring an auto of one of three categories of fuel efficiency in the 12-month period ending February 1980. To minimize violation of the IIA property of the BNML model, this model considering the choice between not acquiring an auto, acquiring an auto of type 1, acquiring an auto of type 2 or acquiring an auto of type 3, the choice has been structured as per equation (9) (see Fig. 2).

![Diagram](attachment:image.png)

**Figure 2: Model Structure**

The estimate of the parameter of EMU can be used to test for suitability of this structure. An important point is that we are modelling the acquisition choice, not the number of household autos choice. That is, we are not explaining different levels of stock but the change in the level of stock (including a composition change associated with a zero change in stock level). As a consequence it is essential that for non-acquirers we obtain details on vehicle(s) that would be acquired had they made an acquisition. Failure to do this has obvious problems (e.g. the EMU variable for the non-acquirers cannot be calculated for the acquisition alternative).
PRELIMINARY RESULTS

The final models estimated on 151 observations are reported in Table 2.

On conventional statistical criteria of goodness-of-fit when the function is non-linear, both models perform well. McFadden (1979) argues that a pseudo-R² between .2 and .4 is an extremely good fit. The alternative-specific constants (ASC) are all statistically non-significant which is a desirable result for a well-specified model.

The EMU variable has an estimated parameter (λ, λ₁) equal to .1664, satisfying the condition for utility maximisation. Its closeness to 0, however, suggests that the decision structure tends to independence rather than simultaneity. This is an important result, suggesting that the sample of households regard the acquisition and type choice decisions as mildly interdependent. The coefficient is significantly different from zero at the 10 per cent level. The finding means that given an improvement in the utility from acquiring an auto of type t, then the likelihood of a household changing from acquiring one auto of type t' to acquiring an auto of type t, will be greater than the likelihood of a household changing from not acquiring an auto to acquiring an auto of type t'. Overall, there will be a greater increase in the expected number of households selecting type t compared to type t'. If a simultaneous structure were used, we would be (incorrectly) ensuring that there will be an equal increase in the choice of all auto types resulting from an improvement in the utility from acquiring an auto of type t.

The variables in the auto-type choice model relate to the physical characteristics of an auto (weight), the performance characteristics (acceleration), usage characteristics (long-distance travel), ownership characteristics (by household or another source, typically a company), and cost characteristics (purchase outlay, expected future operating costs). The diversity of significant explanatory variables confirms our belief that market share cannot be simply forecast by using conventional macro-indicators typical of the aggregate econometric studies (e.g. Chase Econometrics Associates 1974, Wharton Econometric Forecasting Associates Inc 1977). However, it should not be argued that the aggregate models cannot be modified to include such variables and still perform a useful role.

The acceleration rate entered all utility expressions (i.e. as a generic variable), and with a positive significant coefficient. This says that, ceteris paribus, an increase in acceleration performance is likely to increase the probability of selecting all types of vehicles; however, given the current levels of acceleration of each type category, a one per cent increase in acceleration will
### Table 2: A Structured-Logit Model of Auto Type Choice and Auto Acquisition
Fitted by Maximum-Likelihood Method

#### Auto Type-Choice Given Acquisition Level
(alternative 1 = low fuel consumption  
2 = medium fuel consumption  
3 = high fuel consumption)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>estimated coefficient</th>
<th>asymptotic t statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-low fuel consumption (1)</td>
<td>.1392</td>
<td>0.28</td>
</tr>
<tr>
<td>ASC-medium fuel consumption (2)</td>
<td>0.054</td>
<td>0.01</td>
</tr>
<tr>
<td>acceleration rate (0-80 kms/hr, secs)</td>
<td>(-1.70)</td>
<td></td>
</tr>
<tr>
<td>weight #10 (kg) (1, 2, 3)</td>
<td>-1.096</td>
<td>2.01</td>
</tr>
<tr>
<td>retail expenditure value * .01($)</td>
<td>-1.70</td>
<td></td>
</tr>
<tr>
<td>fuel price expectations (%) (2)</td>
<td>0.92</td>
<td>1.85</td>
</tr>
<tr>
<td>regular long distance dummy (2)</td>
<td>0.92</td>
<td>2.03</td>
</tr>
<tr>
<td>source dummy (1 = acquired by non-household sources)</td>
<td>1.3680</td>
<td>2.03</td>
</tr>
</tbody>
</table>

Log-likelihood at convergence: 38.9
adjusted pseudo-R²: .34
proportion successfully predicted: 60%

#### Acquisition-Choice
(alternative 1 = acquired one vehicle,  
2 = acquired zero vehicle)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>estimated coefficient</th>
<th>asymptotic t statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-acquisition (1)</td>
<td>.6469</td>
<td>0.65</td>
</tr>
<tr>
<td>search cost (no. of hours searching)</td>
<td>.9248</td>
<td>1.70</td>
</tr>
<tr>
<td>number of cars per head of licensed drivers (before acquisition but after decision)</td>
<td>-5.374</td>
<td>-4.00</td>
</tr>
<tr>
<td>disposal dummy (1 if dispose a vehicle)</td>
<td>4.1170</td>
<td>5.30</td>
</tr>
<tr>
<td>number of days per week each vehicle is not used #10</td>
<td>.0325</td>
<td>2.40</td>
</tr>
<tr>
<td>EMU (1)</td>
<td>.1664</td>
<td>1.70</td>
</tr>
</tbody>
</table>

Log-likelihood at convergence: -43.64
adjusted pseudo-R²: .55
proportion successfully predicted: 85%
The retail expenditure variable is not statistically significant; hence we cannot meaningfully interpret its magnitude nor sign. The result is arguably due to the definition of the type choice set, purchase price is highly correlated with the dimension of our categories.

The other vehicle characteristic, weight, is significant at the 10 per cent level and has an intuitively plausible sign. It suggests that as the weight of the vehicle increases, ceteris paribus, the likelihood of selecting that vehicle type decreases. A 1 per cent increase in weight will reduce the probability of acquiring a small car by 0.30 per cent, a medium car by 0.46 per cent and a large car by 0.62 per cent. This result indicates more scope to adjust as the size of the vehicle (measured in litres/100 km) increases. A comparison with the USA evidence (e.g. Levine and Train) shows that for Australia there is less scope to use the weight dimension in improving fuel efficiency of the total household fleet.

The retail expenditure variable is not statistically significant; hence we cannot meaningfully interpret its magnitude nor sign. This result is arguably due to the definition of the type choice set, where purchase price is highly correlated with the dimension of our size categories.

A most important variable is fuel price increase expectations, shown in the USA to have a more significant influence on auto-downsizing than the absolute level of petrol prices. The expectations variable in the type-choice model relates to the per cent change in petrol price increases expected between the beginning of 1980 and 1981. The changes between 81 and 82, and 82 and 83 were not statistically significant. Testing for functional form, as absolute or difference variables, did not alter the result. This variable at the 10 per cent level is alternative-specific, and enters the utility expression for medium fuel consumption autos. Like the remaining auto-type variables they are not assigned to alternative 3, only alternative 2 because the large fuel consumption category is predominantly company cars; and these variables are not pertinent in this context. This means that, ceteris paribus, an increase the probability of acquiring a small car by 0.55 per cent, a medium car by 0.74 per cent and a large car by 0.71 per cent. The positive sign contrasts with the Manski and Sherman (1980) finding that for all age groups acceleration was a significantly disliked attribute.

5. Personal discussion with Royce Ginn and Len Sherman revealed that they had anticipated a positive coefficient, and believe that their result is of some concern.

6. A comment is required on the retail expenditure variable, which takes the actual value rather than $0 for company-provided cars. This may, in part, contribute to the statistical non-significance; however, the purpose of adopting this position is to argue that we are picking up the quality dimension, i.e. the more expensive a non-household provided auto, the greater the utility to be gained, and 'even more for nothing'. This is suggested as a reason for an unexpected sign, even though the statistical non-significance makes such a statement highly qualified.
AUTOMOBILE ACQUISITION AND TYPE CHOICE

Expectation of a greater percentage increase in petrol prices will reduce the likelihood of selecting a medium auto. The statistical non-significance of the petrol price expectation variables for later years suggests that although households have well-defined (predictable) expectations extending over a time horizon of one year, expectations beyond this period appear to have insignificant influence on current auto-type decisions. Thus, the perceived risk of not downsizing appears to be very much immediate.

The variable 'regular long distance dummy', reflects the hypothesis that households will prefer at least a medium-sized auto if they travel regular long distances, measured in terms of travelling more than 80 kilometres a day at least once a week. This is entered as a dummy variable (1,0). It is significantly different from zero at the 5 percent level. The sign has no necessary expected direction. For example, a negative sign could be obtained to reflect the hypothesis that people who travel regular long distances prefer a small car so as to economise on fuel. However, for the current data set, such people tend to have medium (and larger) autos, including all the company cars. The comfort dimension (not satisfactorily accounted for in our model) suggests itself as the reasoning behind the result obtained.

We were able to test for the company car effect (the source dummy variable), but it was not statistically significant, probably because it is predominant in the large fuel consumption category and has a small variance for this data set.

We conclude by warning the reader that the results are tentative, but give guidance for further research in this important area.

In the acquisition-choice model, four variables were found to be statistically significant in addition to the EMU variable, already discussed. The search cost variable is designed to account for some of the transactions costs involved in entering the auto market and acquiring an auto. Manski and Sherman (1980), argued strongly for improved measures of transactions costs than a simple dummy variable to distinguish new and used cars (where search costs were assumed higher in the used car market). While this alternative specification is more appealing than the Manski-Sherman formulation, more research is necessary.

7. We investigated the annual kilometres driven in an acquired car, which does not test for the same effect as the regular long trip variable, but does provide a test for effect of time in a vehicle (and hence comfort). It was not statistically significant.

Important areas for further research are the relationship between the way in which kilometres are allocated in vehicle use and the implications (possibly via other variables such as comfort) on vehicle size. The company car cannot be ignored since it clearly has an important interactive influence via the 'costless' effect.
required on the specification of search costs and other transaction costs. The search cost variable enters both the acquisition and nonacquisition utility expressions.

A fundamental problem arose in calculating this variable for the households that did not acquire a vehicle. The survey did not obtain the anticipated hours searching if such a household were to have entered the market and acquired or decided not to acquire an auto. One might hypothesise that their expected search costs are on average higher than the acquirers, giving a major reason for nonacquirers not entering the market. Alternatively, if they entered the market and decided not to acquire their search costs were equivalent to those of the transfer of average search costs based on the levels experienced (new or used), and the presence of one or more company cars in the household fleet. This is likely to produce sample selectivity bias (Heckman 1979). For the acquirers, we assumed a zero search cost for acquisition; that is, they would not have entered the market. In future research we will address the issue of whether a nonacquirer actually entered the market (hence a search cost) or stayed out. This is clearly a fundamental issue. This variable, statistically significant at the 10 percent level, is interpreted as an information-gathering variable, and one can argue that an exhaustive search of the market is likely to yield greater utility (e.g., a suitable vehicle) than a less exhaustive search, thus acquisition is more likely. The positive sign is consistent with this interpretation and hypothesis.

The car stock utilisation variable is significant at the 5 percent level and has a positive sign. This is a somewhat ambiguous variable in terms of its sign. The obtained sign suggests that the ownership of 'excess' stock provides flexibility and hence is a positive contribution to the utility of acquisition (4). Casual empiricism leads us to support the 'on call' interpretation. An alternative hypothesis (consistent with a negative sign) is that as the intended utilisation level increases, it would be expected that there is less likelihood that the intention to acquire another car will be actually carried through.

The final two variables, cars per licensed drivers and the disposal dummy (making disposal exogenous) are both significant at the 5 percent level and of the expected sign.

8. If the acquired vehicle is a replacement, the positive sign could be also interpreted (in a downsizing environment) to suggest that fuel costs may be influencing households in replacing their 'little-used' relatively fuel inefficient vehicle with a relatively fuel efficient vehicle. Alternatively, fuel price increases could lead to acquisition of a 'little-used' vehicle which may not necessarily be a more fuel efficient vehicle; that is, the intention is to use it less.
CONCLUSIONS AND SUGGESTIONS FOR CONTINUING RESEARCH

The empirical application has limitations, some of which have been mentioned. A weakness of this relatively sophisticated model is the consideration of a restrictive acquisition decision - to acquire one vehicle or none, and the very tentative nature of definition of type choice categories. It is straightforward to extend the model (within the same framework) to accommodate all levels of acquisition and disposal of autos. However, a more appropriate model system should include the household decision on the allocation of a budget between expenditure on capital, operating, and maintenance outlays for autos and other commodities, together with the interdependent decisions of vehicle usage, residential and workplace locations, and even mode choice. Other considerations such as vehicle financing and age of the auto may be endogenous (i.e., choices in their own right) or simply exogenous variables in another-choice model.

Even if such a sophisticated set of models were developed, there is a need to move away from the single cross-section approach to a repeated cross-section or panel approach, enabling a true dynamic choice scheme. Important influences such as accumulated experience with the existing or previous stock, search costs, and changes in energy prices over time cannot be well explained in the auto-choice context without a dynamic choice framework. Recent research by Heckman (1981), Johnson and Hensher (1980), LePlastrier (1981), Chamberlain (1980) and Daganzo and Sheffi (1980) extends discrete-choice modelling to a dynamic setting. It is complex, but arguably more realistic; a major source of bias in the past being cross-sectional bias.

The Future of the Car Project in progress at Macquarie University entails monitoring of 2000 households (on an annual basis) in the development of a panel data set in the auto-choice area. This will enable us to develop dynamic empirical models. Until then we refrain from drawing policy implications.
REFERENCES


AUTOMOBILE ACQUISITION AND TYPE CHOICE


HENSHER AND MANEFIELD


McFadden, D., Train, K. and Tye, W (1977) 'An application of diagnostic tests for the independence from irrelevant alternatives property of the multinomial logit model', Transportation Research Record, No. 637.


