The theory and practice of car ownership forecasting *

by
K. J. BUTTON
Department of Economics, University of Loughborough
and
A. D. PEARMAN
School of Economic Studies, University of Leeds

Forecast levels of car ownership play a vital role in the determination of government transportation planning and policy. Both at the urban and inter-urban level, the scale and character of transport provision for the last two decades of this century will be determined to a significant extent by the view taken now about the position of the private car in our society. In the United Kingdom, for example, the British government's recent *Transport Policy* document contained a section specifically devoted to forecasting national car ownership levels. The accuracy of such forecasts has figured importantly in a number of recent controversies in the U.K. surrounding public enquiries into motorway plans. It is clear that well conceived car ownership models are an essential requirement if the uncertainty inherent in the traffic forecasts central to infrastructure investment appraisal is to be minimised.

The first section of this paper is concerned with the theoretical framework within which car ownership decisions may be viewed as being made. It is noticeable, however, that little applied work in car ownership forecasting has endeavoured to incorporate a rigorous theoretical framework of this type. Section 2 discusses a number of the better known types of car ownership forecasting model pointing out some of their strengths and weaknesses and also the frequent inconsistency of their implications, for example, with respect to sensitivity to household income levels, a crucial explanatory variable in many models. Attention is also drawn to the common use of proxy variables and the theoretical and practical difficulties which arise as a result. The final section of the paper summarises the present position in car ownership forecasting and emphasises the need to continue the movement towards behaviourally based models and away from the empiricism which characterised much early work in this area.

THEORETICAL MODELS OF THE CAR OWNERSHIP DECISION

Analysis of the car ownership decision seems to fit most appropriately into the realm of economic theory. In economic terms, the car may be regarded as a consumer durable good. Its purchase represents a medium-term investment of capital which is repayed over the life of the vehicle through the services which its possession provides. However, if the transport planner turns to the literature of economics, either theoretical or applied, for an understanding of this aspect of consumer behaviour, he will get relatively little help. There are a number of reasons for this.

The first problem encountered in specifying an economic model of the demand for cars is to determine who is responsible for the demand. In the main body of neoclassical economic theory it is implicit that the decision maker is the individual consumer. It seems however, that the car ownership decision is far more likely to be a household one, rather than one taken by an individual household member in isolation. The way in which household decisions are reached is a matter which has not received a great deal of attention from economists. The household seems normally to be treated as a single-minded decisionmaking unit, a quasi-individual rather than a collection of individuals. The analytical convenience of such an approach is obvious, given the rich development of theory for individual demand decisions. Its validity is less clear. For some items, where the adult members of the household make similar and approximately equal use of the good concerned, such an aggregation may be justifiable. For something like a car, however, where its use by one member for, say, the journey to work may well deprive other household members of its use altogether, it its harder to justify the ownership decision as a quasi-individual one.

Even if the problem of the nature of the decisionmaking unit is side-stepped, a number of other significant difficulties remain. These involve specifying the good demanded, quantifying the demand and identifying the influences which create the demand and which should, therefore, constitute the explanatory variables in any economic model. Despite the fact that it is normal to discuss the forecasting of demand for cars, this is, to a large extent, a misnoma. The principal demand is for the services of cars, rather than the cars themselves, although, in some kinds of work in this general area, the status symbol aspect of car ownership cannot be overlooked. The importance of the demand for car ownership arises via the demonstrably strong relationship between

* *This paper derives from work undertaken as part of a research project carried out under the auspices of a grant from the Social Science Research Council.*

ownership and use (see Oi and Shuldiner (1962); Deutschman (1967)) and the implications of use levels for infrastructure policy. Thus the models which are of relevance here are predictors of the total stock of cars and not estimators of new sales per unit of time, as in most economic demand studies. This complicates matters considerably. It means that the second-hand market, with all its complexities and problems of data availability, is involved. It means further that, because the car is a durable consumer good, and one which involves a heavy capital outlay, the variables which explain its purchase may well, in part, be transitory and unavailable for observation in the context of recording of ownership some years after the purchase decision has been made. Alternatively, the opposite problem awaits supporters of Friedman's permanent income hypothesis. In this case, a household with an expectation of an increasing income might well take its car ownership decision on the basis of factors as yet unobserved.

Further, the neoclassical theory of demand is most fully developed in terms of continuous functions and variables. It is not well-suited to the discrete 0,1,2, ... car decision that the household makes. The explanatory variables also cause difficulties. Many of them are likely to be outside the set typically embraced by models of demand, factors like price levels and income. There is stronger interdependence with other economic decisions of the household, particularly job choice and household location. Unlike some other markets, the variation in these factors is so great that aggregative studies of market demand are likely to obscure numerous significant features, matters about which one might well wish to have some insight from a policy point of view. For example, if use of cars is likely to become intolerable given present trends, it might be desired to alter household and job location possibilities as a way of cutting car ownership and hence use. Models which aggregate away the effects of locational decisions are thus of limited value. A final difficulty of specification which may be noted is the widely held belief that many car ownership and use decisions are not justified on the basis of the costs involved. Although other factors may explain this phenomenon, inaccurate perception of costs may also have an influence, and one which would be very difficult to estimate.

Although the preceding paragraphs have only sketched in some of the difficulties likely to be encountered in specifying models of car ownership as an economic decision, they do provide some clue as to why very few economists have been trampled to death in the rush to provide such insights to the transport planner. Probably the best known attempt to explain the car ownership decision, starting from a basis in economic theory, is due to Beckmann *et.al.* (1973). In outline, the type of approach they adopt is as follows. Without a car, a household's total utility, U^o, at any time is taken to be dependent on its disposable income, Y, its available leisure time, L, and the travel which it undertakes during the specified period. Without a car, the destinations within reach of the household can be defined as D^o and the total time spent in travel denoted as $\sum_{n \varepsilon D^o} r_n^o T_n^o$, where T_n^o represents the number of trips made to destination n and r_n^o the travel time to reach n. The utility function can then be defined as

$$U^o = U(Y,\ L - \sum_{n \epsilon D^o} r_n^o T_n^o\ ,\ T_n^o) \qquad (1)$$

The purchase of a car will have several effects. It will immediately reduce the income available to buy other goods by an amount, C, equal to the annual cost of car ownership. It will increase the number of destinations accessible to the household from D^o to D^1. Finally, it will affect both the travel time it takes to reach each destination and the number of trips likely to be made to each. The total time spent in travel will become

$$\sum_{n \varepsilon D^1} r_n^1 T_n^1$$

Consequently, the household's utility with a car can be represented as

$$U^1 = U(Y - C,\ L - \sum_{n \epsilon D^1} r_n^1 T_n^1\ ,\ T_n^1) \qquad (2)$$

It now becomes clear that, in general terms, it becomes worthwhile for the household to purchase a car if $U^1 > U^o$. The introduction of additional subscripts enables this simple analysis to be extended to cover decisions about the purchase of a second car, or a better car.

The principal extension of this work is due to Burns *et.al.* (1975). It takes the form of an empirical test of the general theory using data gathered as part of the Detroit Regional Transportation and Land Use Study. In order to operationalise the theory, a specific functional form for the utility functions has to be assumed and also a model of the interaction between individual household members with respect to car ownership and its effect on total household utility. The detailed nature of these assumptions is certainly open to some criticism. So too is the choice of a sixty-minute time parameter for determining D^o and D^1 and the use of an attractiveness index for destinations which is independent of household location. Problems of this nature are, however, virtually inevitable in using existing data sources for the preliminary testing of new theoretical approaches. The overall impression is that the multinomial logit model developed on the basis of this theory is a significant step forward in modelling car ownership behaviour.

A second recent American disaggregated behavioural model of car ownership is due to Lerman and Ben-Akiva (1975). Again, a multinomial logit model is used, but the theoretical foundation is given rather less emphasis, with the principal intention being the incorporation of variables which can reflect the full breadth of potential household behaviour with respect to car ownership decisions. Of particular interest is the incorporation of a variable reflecting the position of each household in its life cycle — a measure of household structure. As in the paper by Burns *et.al.*, there are many points of detail which are contentious, but the general approach is a valuable one, deserving further research input.

One respect in which these recent models are open to criticism, in company with many earlier attempts, is their failure to recognise potential dynamic effects. It seems plausible that at least the rate at which society adopts the motor car, if not ultimate saturation levels of ownership, will be influenced by existing ownership rates. This would certainly be the implication of Tanner's work, described in the following section. The appropriate parameters in applying cross-sectional models, in other words, may well be time-dependent, via the effect of ownership rates in intervening periods on behaviour patterns. This matter is discussed further in Section 2 in the context of Bates' work.

If the principal concern in analysing car ownership is to predict its likely course and the outcome of trying to change that course, a fully dynamic model may, therefore, be the most appropriate. Such a model would recognise explicitly the effect of present car ownership rates on future ones. Clearly, introducing a dynamic dimension further complicates an already difficult modelling task. To demonstrate the kind of model which may be relevant to some circumstances, and to draw out some of its implications, a relatively straightforward expository example will be used.

Consider the obverse of the car ownership model, a model forecasting the number of captive public transport users in an urban area consisting of m zones.[1] Although it would be straightforward to disaggregate the model by household type, consider, for simplicity, a zonal level of analysis. Let R_i^o be the percentage of the residents of zone i who are captive users of public transport. Then the utility accruing to this group through their travel behaviour may, for simplicity, be regarded as equal to the difference between the net benefits gained through the trips undertaken and the cost of those trips. For purposes of demonstration it is assumed that all trips have equal unit benefit value, v, and that the costs incurred are a known multiple, p_i, of the percentage of the population who are captive users. Variations in p_i permit the recognition of potentially different cost structures for producing an appropriate level of service in different zones. Thus

$$U_i^o = v \sum_{j=1}^{m} T_{ij}^o - p_i R_i^o \tag{3}$$

Suppose now that the rate of growth (decline) of the percentage of captive public transport users is some constant multiple, k, of utility.

$$\dot{R}_i^o = kU_i^o = k\left(v \sum_{j=1}^{m} T_{ij}^o - p_i R_i^o\right)$$

Further, suppose that the number of trips, T_{ij}^o, is estimated using an attraction-constrained entropy maximising model so that

$$T_{ij}^o = \sum_j \frac{O_i D_j e^{-BC_{ij}}}{\sum_i O_i e^{-BC_{ij}}} \tag{4}$$

If we assume that the unconstrained number of public transport trips originating in zone i is a simple linear function, $O_i = k^1 R_i^o$, of the percentage of the population captive to public transport, then, by substitution, we have

$$\dot{R}_i^o = kv\sum_j \frac{R_i^o D_j e^{-BC_{ij}}}{\sum_i R_i^o e^{-BC_{ij}}} - kp_i R_i^o \tag{5}$$

$$= \left[kv\sum_j \frac{D_j e^{-BC_{ij}}}{\sum_i R_i^o e^{-BC_{ij}}} - kp_i \right] R_i^o \tag{6}$$

or

$$\dot{R}_i^o = M_i\,(R_1^o \ldots R_m^o)\; R_i^o \tag{7}$$

(6) represents a series of m rather complex simultaneous differential equations. As has been shown, however, by Hirsch and Smale (1974) and by Wilson (1976), some qualitative analysis of this type of equation system is possible.

Following Wilson, the Hirsch and Smale two dimensional analysis generalises to a requirement that three sets of conditions hold with respect to the equation system. If this is the case, then some qualitative conclusions can be reached about the time path of the R_i^o. The conditions which must hold are derived from (6[1]):

(a) $\dfrac{\partial M_i}{\partial R_k^o} < 0, \; i \neq k$

$$\frac{\partial M_i}{\partial R_k^o} = -kv\sum_j \frac{D_j e^{-BC_{ij}} e^{-BC_{ik}}}{\left(\sum_i R_i^o e^{-BC_{ij}}\right)^2} \quad < 0 \text{ if k and v are} > 0$$

This implies that R_i^o decreases if any other R_k^o increases. This might, for example, occur if there are supply inelasticities in the public transport sector.

(b) $\exists$ K such that $M_i \leqq 0$ if $R_i^o > K$

As R_i^o is increased, the summed term in the bracket in (6) decreases. It will ultimately become less than kp_i and so the condition is fulfilled.

(c) $\exists\, J_i$ such that $M_i\,(0,0, \ldots R_i^o, 0 \ldots 0)$ > 0 for $R_o^i < J_i$; < 0 for $R_o^i > J_i$

$$M_i\,(0,0 \ldots R_i^o, 0 \ldots 0) = kv\sum_j \frac{D_j e^{-BC_{ij}}}{R_i^o e^{-BC_{ij}}} - kp_i$$

$$= kv\sum_j \frac{D_j}{R_i^o} - kp_i$$

Clearly, by altering the size of R_i^o, M_i may be made either positive or negative, as required. With these three results established, it is now possible to go ahead and analyse the behaviour of the M_i in phase space $(R_1^o, \ldots R_m^o)$. A detailed analysis is not justified for an exploratory model of the type just outlined, nor is it possible without knowledge of the parameters of the system. However, it is possible to make two points. First, equation sets of this kind can frequently exhibit bifurcation. That is, the equilibrium set of R_i^o to which the system will tend will vary depending upon the initial conditions. An equilibrium with all $R_i^o > 0$ is possible. So is one in which some R_i^o fall to zero. The policy implications are clear. If the current situation corresponds to a point in phase space which leads to an unacceptable equilibrium (say, because it leaves minority groups like the elderly with no public transport provision in certain zones) then an interference with the natural order of things will be required, either to alter the equation system or to change the initial conditions to a set leading to a more acceptable equilibrium. The second point to be made is that, even with the very simplistic model presented here, the complexity of analysis involved in consideration of dynamical systems is clear. Implementation of this type of analysis would not be straightforward. With this in mind, it is interesting now to examine the kind of models which economists and transport planners have tried to implement and to compare them both with each other and with theoretical ideals.

EXISTING MODELS OF CAR OWNERSHIP

In practice there has been little attempt to employ economic models of car ownership for forecasting purposes. Rather, there has been an emphasis on straightforward empirical work, concentrating on techniques which offer a reasonably good statistical explanation of past and current levels of car ownership to predict future growth. Broadly these methods can be divided into two main categories, those employing some form of extrapolation procedure to trace out the past growth in vehicle ownership and to extend this trend into the future and those adopting crude econometric methods, usually, but not exclusively, employing cross-sectional data to determine certain statistical parameters which are then assumed invariant with respect to time. The former method is most widely used at the macro level for national forecasting, while the latter tends to be favoured at the more micro level either for regional or local forecasting.

The extrapolation techniques generally assume that car ownership follows some form of sigmoid growth path through time until ultimate saturation level of ownership has been attained. This is, for example, the standard procedure employed by the Transport and Road Research Laboratory in the U.K. (Tanner (1974)) to draw up their national forecasts. The sigmoid growth path is described by a logistic curve with an exogenously determined saturation level used as an asymptote. Although slightly different methods were used to determine the saturation levels, the logistic curve fitting technique has also been used on Dutch car market (Bos (1970)) and on the U.S. car market (Whorf (1975)). This type of procedure is useful if the only information required is a rough estimate of the car stock at some future date. It has some basic justification in that income is likely to be an important explanatory variable of the car ownership level and that there has been a long term trend for income to increase through time. By relying upon time as the independent variable the logistic curve fitting procedure implicitly circumvents the problem of trying to forecast future income levels.

The detailed method of logistic curve fitting, and in particular, the T.R.R.L. approach, has been subjected to serious criticism in recent years. Although the extrapolation approach removes the need to predict future values for a set of explanatory variables it still requires an estimate of the ultimate saturation level of vehicle ownership, and this is in many ways equally difficult. In the past the T.R.R.L. have attempted to devise an objective statistical method for arriving at this saturation level. They employed cross-sectional information on the annual rate of change in car ownership in each county as the dependent variable in a regression run against the actual ownership level. The ownership level at which the rate of change becomes zero is then used as the saturation level in the logistic curve fitting exercise. Criticisms of this procedure have ranged from the unjustified nature of the orthoganal regression used to determine the saturation level to the T.R.R.L.'s selective use of data in arriving at their final result (Adams (1975)). Even empirically the procedure was unsatisfactory and it became apparent in the early 1970s that the saturation level derived depends crucially upon the year for which the cross-sectional data is taken. There is a tendency for the saturation level to increase over time if sucsessive annual calculations are performed. More recent work at the T.R.R.L. has resulted in the adoption of a much more flexible approach to estimating the saturation level involving a consideration of the eventual number of drivers in the population and a review of trends in other countries. However, it still does not have any underlying theoretical basis to describe the forces working towards this saturation level nor an economic theory of why such a saturation level is inevitable. If we look at the work undertaken in the U.S. by Whorf, we find a similar weakness, he simply runs a series of regressions corresponding to the logistic curve using arbitrary values for a saturation level. The model offering the highest $\bar{R}^2$ is then selected. There is no theoretical justification given for the saturation level finally arrived at, it simply emerges from the data.

For the local planner the logistic curve fitting technique has two fundamental weaknesses. Firstly, the urban transport planner is only indirectly interested in car ownership as such, he is primarily concerned with car use. Car ownership is, therefore, used as an input into a further series of models which forecast travel patterns in the urban area. Aggregate numbers of vehicles are not the important consideration here, travel habits are more strongly influenced by the number of vehicles available to each individual household. Unfortunately, the logistic procedure does not offer information on the number of no car, one car, and two plus car households, only the average ownership level. Secondly, any form of extrapolation implicitly assumes that underlying influences do not alter the *ceteris paribus* conditions — but this is unrealistic at the disaggregated level. Indeed, the transport planner, by modifying the local transport system, will himself disturb these *ceteris paribus* conditions.

In consequence, local car ownership forecasts are distinguished by their use of cross-sectional data in combination with very simply statistical models. A common procedure is to employ category analysis in which a multi-dimensional matrix is defined with each dimension representing an explanatory factor described in discrete categories, for example, households in an area may be categorised by their income group, the number of residents and their social status (Mogridge and Eldridge (1970)). By assuming the average rate of car ownership for each category does not change over time and by forecasting the future number of households falling into each category it is then possible to predict future car ownership rates. The limitations of this approach are clear. It assumes that the average ownership rate in each category is invariant with respect to time, hence it is essentially a static model being used for a dynamic purpose. In addition, there are statistical problems involved in establishing the contribution of each variable to the ownership rate which can only be resolved by messy analysis of variance tests. Finally, category analysis involves expressing certain naturally continuous variables, such as income, in a discrete form which can lead to distortions in the forecasts made.

An alternative to the category analysis approach is the use of multivariate regression techniques. Again the tendency is to employ cross-sectional data from a transportation survey but regression procedures enable continuous variables to be incorporated in their natural form whilst discrete variables can be represented as dichotomous dummy variables taking the values of 0 or 1. Ideally the regressions are performed at the household level to minimise the variation within observations generally encountered using zonal data. This does present certain statistical problems due to the limited range of values the dependent variable can take (ie., 0, 1 or 2+). Only by sacrificing statistical simplicity can this problem be resolved. One method of circumventing the difficulty is to transform the dependent variable into a probability of car ownership, an acceptable way of doing this is to adopt logit analysis. This statistical manoeuvre is not without its own shortcomings, however, and the conventional standard error tests on independent variables cease to be appropriate in the logit formulation.

The logit model has a particular advantage over other forms of micro ownership forecasting in that it does offer a meaningful economic explanation of causal influences resulting in car ownership growing through time (Bates (1971)). If income is used as the sole explanatory variable in the specification

$$P_o = \frac{c}{y^{-b} + c} \tag{7}$$

(where:- P_0 is the probability of not owning a car

Y is income

and b and c are parameters to be estimated) we can see that as Y^{-b} approaches zero, P_0 will approach unity, but as Y^{-b} rises so P_0 must fall. Transforming this into a logit gives

$$\ln\left(\frac{P_o}{1-P_o}\right) = a + b \log Y \tag{8}$$

Now we can see that the value of b represents the income elasticity of the odds in favour of a family not owning a car. (Theil (1971)). We can also see that $-\left(\frac{a}{b}\right)$ is the natural logarithm of that income at which the household is as likely or not to own a car. If we examine any changes in the model parameters over a series of cross-sections it is possible to separate the effects of the car market as a whole from other forms of expenditure (if a is held constant, but b is observed to rise over time then there has been a general shift in favour of the car market *vis a vis* other markets) and also to explore the interaction of price and the age of the vehicle stock within the market (with $-\left(\frac{a}{b}\right)$and held constant a fall in the value of b over time would suggest the spread of ownership was becoming less unequal indicating a probable fall in the price of older cars relative to newer ones).

Subsequent work using this approach has seen the range of explanatory variables employed expand (Fairhurst (1975)) but its use as a forecasting tool is still very limited. In order to produce reliable forecasts from cross-sectional data, the conventional regression analysis approach assumes the model parameters represent long-term elasticities which do not change over time. By its very nature, however, the type of model set out above makes an entirely different assumption, namely that these parameters can vary but that this variation can be given a sensible economic interpretation. To forecast car ownership accurately using such a model requires, therefore, not only estimates of the future magnitudes of a series of independent variables but also some knowledge of how their associated coefficients are going to behave.

Reliance upon empiricism in car ownership forecasting has resulted in two particularly unsatisfactory consequences. Firstly, there is seemingly little consensus about the importance of fundamental economic variables in car ownership regression models. Let us take income as an example. At least two studies (Fishwick (1972) and Bos (1970)) have questioned its importance as an explanatory variable on the grounds of statistical significance. As we observe from Table I the other major studies which do include an estimate of the income elasticity of car ownership seem unable to reach an approximate consensus upon the value of such a parameter. One can perhaps explain away some of the variation in the latter in terms of the type of data employed (time series, cross-sectional, pooled), the level of aggregation adopted (individuals, households or regions), the specification of the model (log-linear, linear or semi-logarithmic) and the other variables included in the regressions but these factors only reinforce the fundamental criticism rather than weaken it. The specifications were accepted on their statistical merits rather than their theoretical soundness and the very multiplicity of results simply illustrates the inadequate theoretical foundations upon which these empirical models are based. Certainly one cannot expect that every attempt to calibrate a model of car ownership decisions will yield identical parameters, but a more solid and rigorous approach to the underlying causal relationships would result in much more consistency and a much narrower dispersion of the parameters obtained.

Secondly, many studies rely for their explanatory powers upon variables which have no readily, identifiable economic rationale. In some cases attempts have been made to justify the use of such variables by adopting the argument that they are acting as proxies or surrogates for influences which are either not immediately quantifiable or for which there is no readily available data. These 'artificial' variables are of three types. Firstly, there are actual variables which may add considerably to the explanatory power of a regression model but which have only an indirect claim to being called explanatory variables. Spatial parameters such as residential density or the percentage of a region's population living in conurbations fall into this category. *Ex post* justifications for their inclusion usually mention their role as proxy for local public transport quality or for accessibility more generally defined. An example of their importance can be seen in the following equation based upon pooled data from English and Welsh standard regions for 1965-72:

$$C = -0.762 + 0.0008Y^{*} - 0.018PD^{*} + 0.027U^{*} + 3.317S^{*} + 0.008H \qquad \bar{R}^2 = 0.7226 \quad \ldots (9)$$

Where C = Cars per household
Y = Household income net of direct taxes & other deductions
PD = Population density
U = Level of unemployment (%)
B = Employment in basic industries (% of total labour force)
S = Social Economic Group (% in SEGs 1, 2, 3, 4, and 13)
H = Household size
* indicate the variables significant at the 99% level.

One may reject the model in detail because of certain ambiguities in some of the coefficients (e.g., car ownership appears to rise with the level of unemployment) but the P.D. variable is likely to be retained because of its high level of significance and because it would be argued that one would expect car ownership to be lower in regions which are densely populated and likely to have adequate public transport. For forecasting this is not very helpful, however, because the planner is likely to alter the relationship between density and

Table 1

Study	Period Covered by Study	Data Source	Income Elasticity
Evans (a)	1948-64	United States	2.2
Suits (a)	1929-42 & 1948-56	United States	4.2(d)
Cramer (b)	1953	Great Britain	0.69
Kain (b)	1953	United States	0.17(e)
Bennett (b)	1955	United States	1.6
	1956	United States	1.53
	1957	United States	1.67
O'Herlihy (a)	1948-61	Great Britain	1.73-2.48(f)
Kain & Beesley (b)	1960	Leeds	0.72
Smith (b)	1968	United States	0.42-2.04(g)
Sleeman (b)	1968	Less-urbanised British regions	1.76
		More-urbanised British regions	2.89
Buxton & Rhys (b)	1968	English and Welsh regions	1.3
		English regions	1.54
		Less-urbanised English & Welsh counties	0.58
		More-urbanised English & Welsh counties	3.06
	1969	English and Welsh Regions	1.33
		English Regions	2.56
		Less-urbanised English & Welsh counties	0.52
		More-urbanised English & Welsh counties	2.92
Shepherd (a)	1955-71	Sydney	0.347(e)
		Perth	1.032(e)
Pearman & Button (c)	1965-72	English regions	0.3-0.7(h)

Notes
(a) Short run elasticities from time series
(b) Long run elasticities from cross sections
(c) Pooled cross-section and time series data
(d) Relates only to new cars
(e) Calculated from simultaneous models with car ownership treated as an endogenous variable
(f) Sensitive to the rate of vehicle appreciation assumed
(g) Variation between quantities for 1968
(h) Sensitive to the definition of income

Sources: - M. K. Evans; **Macroeconomic Activity, Theory, Forecasting & Control** (Harper & Row) 1969. D. Suits; The Demand for automobiles in the U.S.A. 1929-56 **Review of economics & Statistics** Vol. 40 1958 pp. 273-280. J. S. Cramer; **The Ownership of Major Consumer Durables,** University of Cambridge, Dept. of Applied Economics Monograph 7, 1962. J. F. Kain; A contribution to the urban transportation debate: an econometric model of urban residential and travel behaviour. **Review of Economics & Statistics** Vol. 47, 1964, pp. 55-64. W. B. Bennett; Cross-section studies of the consumption of automobiles in the United States **American Economic Review** Vol. 57, 1967, pp. 841-850. C. St. J. O'Herlihy; Demand for cars in Great Britain **Applied Statistics,** Vol. 14, 1967, pp. 162-195. J. F. Kain & M. E. Beesley; Forecasting car ownership and use **Urban Studies,** Vol. 2, 1965, pp. 163-185. R. P. Smith; **consumer Demand for Cars in the U.S.A.** (University of Cambridge, Dept. of Applied Economics Occasional Paper 44) 1975. J. F. Sleeman; A New look at the distribution of private cars in Britain **Scottish Journal of Political Economy,** Vol. 16, 1969, pp. 306-318; M. J. Buxton & D. G. Rhys; The demand for car ownership: a note **Scottish Journal of Political Economy,** Vol. 19, 1972, pp. 175-181; L. E. Shepherd; An econometric approach to the demand for urban passenger transport **Australian Road Research Board Proceedings,** Vol. 6, 1972, pp. 214-245. A. D. Pearman & K. J. Button; Regional variations in car ownership, **Applied Economics** Vol. 8, 1976, pp. 231-233.

public transport as part of the planning exercise.

A further 'artificial' variable is a time trend. This, it is argued, indicates the autonomous growth in car ownership which cannot be explained in terms of economic influences. We can introduce this into the above model very easily:-

$$C = -1.7367 + 0.0003Y^* - 0.066PD - 0.012U + 2.112B^* + 2.790S^* + 0.500H^* + 0.038T^* \qquad \bar{R}^2 = 0.8605 \quad \ldots (10)$$

where T = A time trend with 1965 = 1

The introduction of T improves the explanatory power of the model in terms of $\bar{R}^2$ and also results in some of the traditional economic variables, notably unemployment, reverting to a coefficient exhibiting the sign one would anticipate. The limitation for forecasting of this approach is that the time trend must be assumed to continue unchanged in the future. In many ways this is an identical assumption to that underlying the extrapolation techniques discussed above and is open to similar criticisms.

Finally, artificial variables can be in the form of 'dummies' which take the value 1 if the region falls into some specified category and a zero otherwise. Figures I(a) and I(b) show the regional growth paths of car ownership per household and per person. Three groupings emerge: (a) North-West, Yorkshire and Humberside and North, (b) Greater London, East Midlands and West Midlands and (c) South East, South West and East Anglia. For statistical reasons we only use dummy variables for the last two groups. The following regression is obtained

$$C = 0.567 + 0.0003Y^* - 0.011PD^* + 0.025U^* - 0.250B + 0.161S - 0.167H + 0.259D_1^* + 0.155D_2^* \qquad \bar{R}^2 = 0.8957 \quad \ldots . (11)$$

where $D_1 = \begin{cases} 1 \text{ if the observation is in regional group (c)} \\ 0 \text{ otherwise} \end{cases}$

$D_2 = \begin{cases} 1 \text{ if the observation is in regional group (b)} \\ 0 \text{ otherwise} \end{cases}$

In purely mechanical terms this equation form is a considerable advance on those preceding it; the $\bar{R}^2$ value is higher, there is less multicollinearity, autocorrelation is considerably reduced and the constant term has a more reasonable positive value. For forecasting purposes, however, one must assume that the parallel growth trends of the three regional groupings will continue with neither any convergence nor divergence and that individual reigons will continue to exhibit the same trends as their parent group. Such assumptions are unlikely to be valid in the longer term but without any theoretical knowledge of why the initial groups occur it is impossible to base forecasts on any other footing.

CONCLUSION

This paper has looked at car ownership forecasting both as a theoretical exercise and as a practical one. Apart from some potential use for time trend extrapolation models at the highest levels of aggregation, the conclusion reached is that future modelling effort should be concentrated on the development of disaggregated behavioural models with a foundation in economic theory. It is possible to have much more confidence in forecasts derived from an inductive approach of this nature than from the deductive models which have commonly been used in the past. Within this latter set of models there is not only strong empirical evidence of inconsistency, but also the ever present danger, when proxy variables are employed, that statistically significant parameters will be justified on an *ex post* rather than *a priori* basis.

The whole concept of developing and calibrating models using existing experience and existing data as a foundation for long-term forecasts is

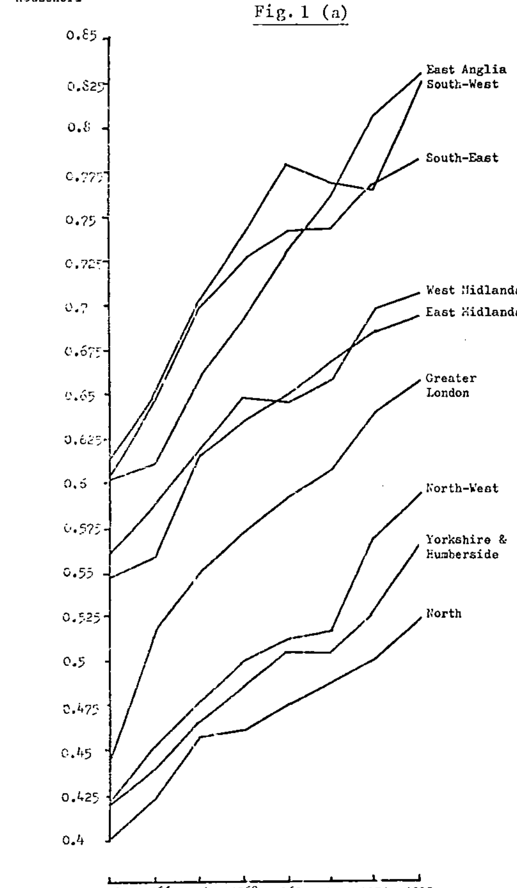
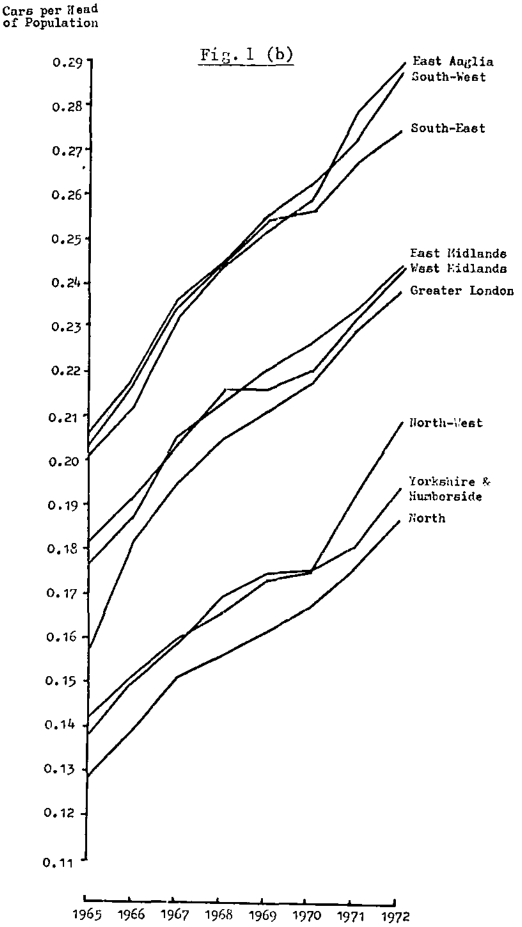


Fig. 1 (a)

Cars per Household in English Regions 1965-72

fraught with difficulties. One whose importance exceeds the brief discussion given to it here is potential dynamic variation in the parameters of a model as a result of levels attained by the dependent variable which are outside current experience. The possibility of a ratchet effect, like that discussed by Duesenberry in consumer theory, should not be overlooked. Data sources with reflect a predominance of decisions to abandon car use are rare. There is some danger in implying that car ownership models calibrated in an environment, say, of generally decreasing public transport provision can automatically be used to predict the effect of a reversal of this trend. New car ownership may be more reliable in this respect, but even here there is the danger of emulation.



Cars per Head in English Regions 1965-72

Problems of this nature are of long standing. To overcome them is difficult and may be impossible. It is, however, wise to be aware of the dangers. In such circumstances, predictions should at least be tempered with some measure of qualification. It is a well known unwritten law that all economists must quote at least once from J. M. Keynes in every paper they write. We do so in conclusion by way of stating our attitude to what car ownership models should seek to achieve.

"The object of a model is to segregate the semi-permanent or relatively constant factors from those which are transitory or fluctuating so as to develop a logical way of thinking about the latter, and of understanding the time sequences to which they give rise in particular cases."

REFERENCES

J. Adams (1974): Saturation planning. **Town and Country Planning,** December pp. 550-554

J. J. Bates (1971): A hard look at car ownership modelling, **Mathematical Advisory Unit Note 216** (Department of the Environment)

M. J. Beckmann, R. L. Gustafson and T. E. Golob (1973): Locational factors in automobile ownership decisions. **Annals of Regional Science** December pp. 1-12

G. G. J. Bos (1970): **A Logistic Approach to the Demand for Private Cars** (University of Tilburg Press)

L. D. Burns, T. F. Golob and G. C. Nicolaidis (1975): Theory of urban-household automobile-ownership decisions. **Transportation Research Record 569** pp. 56-75.

H. D. Deutschman (1967): Auto-ownership revisited: a review of methods used in estimating and distributing auto-ownership, **Highway Research Record 205,** pp. 31-49.

M. H. Fairhurst (1975): The influence of public transport on car ownership in London, **Journal of Transport Economics and Policy,** September pp. 193-208.

M. W. Hirsch and S. Smale (1974): **Differential Equations, Dynamical Systems and Linear Algebra** (Academic Press).

S. R. Lerner and M. Ben-Akiva (1975): Disaggregate behavioural model of automobile ownership. **Transport Research Record 569,** pp. 34-55.

M. Mogridge and D. Eldridge (1970): **Car Ownership in London** (G. L. C. Research Memorandum 387).

W. Y. Oi and P. W. Shuldiner (1962): **An Analysis of Urban Travel Demands** (Northwestern University Press).

J. C. Tanner (1974): Forecasts of vehicles and traffic in Great Britain. **Transport and Road Research Laboratory Report 650.**

H. Theil (1971): **Principles of Econometrics** (North Holland).

R. P. Whorf (1973): Models of automobile ownership, **Proceedings of International Conference on Transportation Research** (College of Europe/Transportation Research Forum).

A. G. Wilson (1976): Towards models of the evolution and genesis of urban structure, **University of Leeds, School of Geography Working Paper 166.**

FOOTNOTE

1. The model developed here draws significantly from the shopping model analysis contained in Wilson (1976).