

TOPIC 15 TRAVEL CHOICE AND DEMAND MODELLING

CONSTRAINED DIFFUSION MODELS FOR THE PREDICTION OF MULTI-CLASS MOTOR VEHICLE OWNERSHIP

CHI-KANG LEE

Department of Transportation and Communication Management National Cheng Kung University Tainan 701 TAIWAN

MING-SHONG SHIAW

Department of Transportation and Communication Management National Cheng Kung University Tainan 701 TAIWAN

Abstract

The ownership structure of motor vehicles in Taiwan is quite different from that in the developed countries. The paper reports a process for building and testing the diffusion models on the basis of traditional logistic curves, including the specification and estimation of saturation levels. As reported in the paper, the newly developed models give promising results in an empirical study with Taiwan's licensing data of car and motorcycle ownership.

INTRODUCTION

Time-series models of logistic curves have been used to estimate the growth trends of car ownership in practice, eg Tanner (1977). As discussed in Mogridge (1989) and in Jansson (1989), such a model provides the long-term dynamics and diffusion process of car ownership. However, as described by Khan and Willumsen (1986), there are some basic differences in the need and use of car ownership models between transportation studies in developed and developing countries. One example is that car and motorcycle in Taiwan are two popular and competitive modes for people to own and to use. Hence, the car ownership models in use in developed countries have to be generalized for the case of three and more alternatives. In the paper, the concept of innovation diffusion provides the basis to build generalized logistic curves. Moreover, various studies used different saturation levels, eg Mogridge (1967), Tanner (1977), and Mogridge (1983), and the growth pattern of a logistic curve is usually sensitive to its saturation level. Hence, the specification of a saturation level in building and estimating a diffusion model is explicitly considered in the paper. Finally, the paper reports the empirical and comparative results of traditional and generalized logistic curves, with and without saturation constraints, for the prediction of car and motorcycle ownership in Taiwan.

MODEL BUILDING

Unconstrained diffusion models

Concepts of temporal and spatial diffusion have been applied to the analysis of geographic, demographic, ecological, economic, and marketing changes; eg refer to Rogers (1983). For a twoalternative logistic curve, Casetti (1969) had shown that it is possible to generate the S shape change on the basis of innovation diffusion. For the car ownership problem, assume that a person is an adopter or a non-adopter of car. At time period t, let $Y_1(t)$ be the absolute number of adopters of car, and $Y_3(t)$ the number of non-adopters. At time t, the maximal amount of personal contacts between adopters and non-adopters can be represented as $AY_1(t)Y_3(t)$, where $0 \leq A \leq 1$. Factor A represents the intensity of socio-economic activities in the study area. If an actual contact causes a change on an adopter's or a non-adopter's choice, it is an effective contact. Let B_1 and B_3 be the proportion of effective contacts respectively for the alternative of car and the alternative of no car. It follows that the change of the adoption level of the car is written as the equation (1):

$$
dY_1(t)/dt = (B_1 - B_3) AY_1(t)Y_3(t)
$$
 (1)

If P is the population of the area, the differential equation (1) can be rewritten as the differential equation (2):

$$
dy_1(t)/dt = a_{13} y_1(t) y_3(t)
$$
 (2)

where $a_1 = B_1AP$, $a_3 = B_3AP$, $a_{13} = a_{1.43}$, y_1 (t)=Y₁ (t)/P, and $y_3(t) = Y_3$ (t)/P. From (2), it is clear that the change of the ownership level of car, $dy_1(t)/dt$, is dependent on the interaction between the two people groups, $y_1(t)$ and $y_3(t)$, where a_{13} is an interaction factor. Moreover, the equation (2) is listed and described graphically in Table 1 as the diffusion structure of model I. By the same process, a differential equation can be obtained and described for the change of motorcycle ownership level:

$$
dy_2(t)/dt = a_{23} y_2(t) y_3(t)
$$
 (3)

where $y_2(t)$ is the ownership level of motorcycle and $y_3(t)$ the proportion of people with no motorcycle.

Table 1 Diffusion structures and differential equation systems

If there are three competitive alternatives (car, motorcycle, and no motor vehicle), and three groups of people. Let $y_1(t)$ = the ownership rate of car, $y_2(t)$ = the ownership rate of motorcycle, and $y_3(t)$ = the proportion of non-motorized people. The interaction relationship among the three groups of people is defined in Table 1 as the diffusion structure of model II. Then, its diffusion model can be readily obtained as the differential equation system of model II listed in Table 1. Such a model was also discussed in Sonis (1986) for the dynamic choice of alternatives.

Constrained diffusion models

Because of the regulation of motor vehicle ownership and the limit of people's ability to buy a motor vehicle, some people are not able to own any motor vehicle. As the diffusion structure of model III shown in Table 2, car ownership and motorcycle ownership are considered independently. For the car ownership problem, s₃₁ represents the proportion of non-motorized people who are inactive to choose a car. By the same process used in the previous section, the differential equation for the change of car ownership is:

$$
dy_1(t)/dt = a_{13} y_1(t)[y_3(t) - s_{31}]
$$
\n(4)

However, the saturation level for car ownership is $1 - s_{31}$, if $a_{13} > 0$. Similar consideration can be used for the motorcycle ownership problem. Then, we have the change of motorcycle ownership as:

$$
dy_2(t)/dt = a_{23} y_2(t)[y_3(t) - s_{32}]
$$
\n(5)

where s_{32} represents the proportion of non-motorized people who are inactive to choose a motorcycle.

Table 2 Constrained diffusion structures and differential equation systems

For the case of three alternatives and three groups of people with one saturation constraint, as the diffusion structure of model IV shown in Table 2, the car ownership and motorcycle ownership are considered simultaneously. The constraint in the diffusion model is that some non-motorized people are inactive to choose any motor vehicle. s₃ denotes the proportion of inactive nonmotorized people, and $1-\text{ }s_3$ is the saturation level for the ownership of motor vehicle. With s₃ and the three-alternative diffusion structure, we have obtained a differential equation system for model IV and listed it in Table 2.

For the problem of car and motorcycle ownership, model V considers three types of inactive groups as diffusion constraints. Let s_{31} = the proportion of non-motorized people inactive to choose car, s_{32} = the proportion of non-motorized people inactive to choose motorcycle, and s_{21} = the proportion of motorcycle people inactive to choose car. Because the price of car, P_c , is usually higher than that of motorcycle, \hat{P}_{m} , s₃₁ is in general larger than s₃₂. With the three-constraint and three-alternative diffusion structure, we have obtained a differential equation system for model V and listed it in Table 2.

Logistic curves

The solutions of the differential equations (2) and (3) are listed in Table 3 as the logistic curves for model I. It is well known that $y_1(t)$ of model I is a traditional logistic curve and it has a S shape of growth if a_{13} >0. Hence, a_{13} represents the net propensity for people to choose car instead of its competitive alternative. For model II, given $y_1(t)+y_2(t)+y_3(t)=1$ or $dy_1(t)/dt+dy_2(t)/dt+$ $dy_3(t)/dt=0$, we have $a_{ij}=-a_{ji}$. Then, the solution for the differential equation system of model II is obtained and listed in Table 3 as a generalized logistic curve for three alternatives.

The differential equation systems of model III and model IV can be solved to have their corresponding constrained logistic curves, which are listed in Table 3. However, there is no closed form for the solution of the differential equation system of model V, so that there is no closed form for the logistic curve for model V. However, given s_{ij} , a_{ij} can be estimated; then, $y_i(t)$ can be computed using the Runge-Kutta method (Andrews 1991).

MODEL ESTIMATION

Linearization

In order to implement the developed logistic curves, linearization is the first step for model estimation. The linearized model I is:

$$
\ln[y_1(t)/(1-y_1(t))] = a_{13}t+c_{13}, \text{ and}
$$

\n
$$
\ln[y_2(t)/(1-y_2(t))] = a_{23}t+c_{23}
$$
 (6)

The parameters a_{13} and c_{13} , and a_{23} and c_{23} can be readily estimated with the linearized equations by regression techniques. Moreover, because $dy_1(t)/dt + dy_2(t)/dt + dy_3(t)/dt = 0$, the reduced and linearized three-alternative diffusion model for model II is a system of two simultaneous equations:

$$
ln[y1(t)/y3(t)] = a13 t + c13, and
$$

$$
ln[y2(t)/y3(t)] = a23 t + c23
$$
 (7)

Thus, the estimation method for simultaneous equations, eg the seemingly unrelated regression method, can be used to calibrate the model parameters a_{13} , a_{23} , c_{13} , and c_{23} . By the same way, each of the constrained diffusion models has a corresponding linearized model, except model V. The linearized models of the logistic curves developed in the previous section are summarized in Table 3.

Table 3 Logistic curves and their corresponding linearized models

À

Since (1) $a_{ji}=0$, (2) $a_{ij}=a_j-a_j$, (3) $a_{ji}=-a_{ji}$, and (4) Σi (d y_i(t)/dt)=0, the number of unknown independent parameters in model V is reduced to two. Hence, a_{13} and a_{23} can be estimated by the following differential equation system:

$$
a_{13} = \{ [(dy_1/dt)(y_2)(y_3 - s_{32})] - [(dy_3/dt)(y_1)(y_2 - s_{21})] \} / \Omega
$$
 (8)

$$
a_{23} = \{ [(dy_2/dt)(y_1)(y_3 - s_{31})] - [(dy_3/dt)(y_1)(y_2 - s_{21})] \} / \Omega
$$
 (9)

where

 $\Omega = (y_1) (y_1) (y_2 - s_{21}) (y_3 - s_{31}) + (y_1)(y_2 - s_{21})(y_2)(y_3 - s_{32}) + (y_1)(y_2)(y_3 - s_{31})(y_3 - s_{32})$

Then, a_{31} , a_{32} , a_{12} , and a_{21} can be obtained by $a_{ii} = a_{ik} - a_{jk}$.

For a constrained diffusion model, the estimation of s_{ij} and the estimation of a_{ij} have to be done sequentially. In this study, a_{ij} is calibrated under many fixed values of s_{ij} . Then, a set of a_{ij} and s_{ij} is chosen on the basis of the statistical performance of its associated model. Therefore, a good experiment design for the saturation level s_{ij} is important for the estimation of a constrained diffusion model.

Diffusion speed

Two kinds of parameters have to be estimated for a diffusion model: a_{ij} and s_{ji} . The sign of model parameter a_{ji} represents the relative competition ability of alternative i to that of alternative j. For example in model I, the car ownership is increasing if a_{13} >0. Moreover, the size of a_{jj} indicates the relative diffusion speed of alternative i. For example in model I, the car ownership will approach its limit very fast if a_{13} is big.

Diffusion speed parameter a_{ji} can be specified as a function of socio-economic variables, so that the frequency of effective socio-economic contacts and the competitive attributes of each alternative can be explicitly represented. However, if it is difficult to find an appropriate function specification and/or a set of socio-economic data, a_{ij} can be simply defined as a function of time variable t or a constant.

Saturation level

The definition of saturation levels has an important effect on model building and model estimation. From the viewpoint of model structure, one example is that the difference between model IV and model V comes from their different definitions of saturation levels. From the viewpoint of model calibration, as discussed in earlier, the linearized models can be estimated by regression techniques if their saturation levels are given.

As discussed earlier, three types of diffusion constraints, s_{31} , s_{32} , and s_{21} , are considered in the model building process. Two kinds of factors have an effect on these saturation levels. First, very young people can not own any motor vehicle. Hence, the proportion of people under 18, UNADU, is considered as an explanatory variable for s_{31} and s_{32} . Secondly, an individual's choice set of motor vehicle ownership is dependent on his ability to pay. For example, a part of non-motorized adult people, $y_3(t)$ – UNADU(t), can not afford a car so that they are inactive in the diffusion process between $y_3(t)$ and $y_1(t)$. Therefore, we can write s₃₁ as:

$$
s_{31}(t) = [1 - \exp(k_1 \times P_c(t) / N I(t))] [y_3(t) - UNADU(t)] + UNADU(t)
$$
 (10)

It is clear that the proportion of inactive non-motorized people for buying a car will increase with an increase in car price (P_c) and/or a decrease in income level (NI). By the same concept, all saturation levels in constrained diffusion models, as shown in Table 4, are defined and specified as functions of socio-economic variables.

Table 4 Saturation levels for the constrained diffusion models

Notes

- 1. P_c and P_m are the prices of car and motorcycle respectively (millions NT\$).
- 2. UNADU is the proportion of people under 18 (%).
- 3. NI is per capita income (millions NT\$).

EMPIRICAL RESULTS

Data

The models described in the second section have been tested in a case study with the licensing data of car and motorcycle in Taiwan. The time-series licensing data from 1968 to 1990 is used for model calibration, and the latest three-year data is used to check the short-term prediction ability of the calibrated models. Moreover, some socio-economic data of the same time period are collected for the calibration of diffusion speed and saturation levels.

Model

The empirical results of the five types of models are listed in Table 5. The diffusion speed parameter a_{ij} in a unconstrained model, model I or model II, is respectively specified as a constant and a function of socio-economic variables. A lot of effort have been spent in the selection of socio-economic variables and function specifications for the diffusion speed a_{ij}. However, only four variables: NI, P_c , P_m , and UNADU, are chosen to describe the interaction relationship between different groups of people for their motor vehicle ownership behavior. As discussed earlier, these variables are also used to describe the change of saturation levels. Each parameter in model I(B) and model II(B), as shown in Table 5, gives a correct sign for its associated socioeconomic variable. The size of each parameter of model I(B) and model II(B) shows its relative effect on diffusion speed. For example in model $II(B)$, it is clear that an one dollar increase in P_c has less effect than a dollar increase in P_m on the diffusion of people from motorized to nonmotorized groups.

The diffusion speed parameter a_{ij} in a constrained diffusion model, model III, model IV, or model V, is specified as a constant or a function of time variable, because the saturation level parameter s_{ij} in such a model has already been specified as a function of socio-economic variables. For model V, many sets of values for k_1, k_2 , and k₃ have been tried to compute best estimates for a_{ij}

PREDICTION OF MULTI-CLASS VEHICLE OWNERSHIP LEE & SHIAW

After that, the parameters in the functions of saturation levels are selected as $k_1=1.2$, $k_2=1.8$, and k_3 =0.3. Since $k_2 > k_1$ and $P_c > P_m$, as the diffusion structure of model V shown in Table 1, the proportion of inactive non-motorized people for car is in general larger than that for motorcycle.

Table 5 Estimation results of the linearized models

Notes:

- 1. The value in a parenthesis is the t statistic.
-
-
- 2. DW is the Durbin-Watson statistic.
3. R² is the adjusted R-square.
4. Ni, P_c, P_m, and UNADU are as defined in Table 4.

TOPIC 15 TRAVEL CHOICE AND DEMAND MODELLING

With the consideration of the adjusted R-square for a model's replicative ability, the Durbin-Watson statistic for an autocorrelation test, and the t-statistic for the significance of each parameter, as shown in Table 5, the empirical results show fairly good statistical performance for most models. Hence, it, is hard to compare the effect of different model structures only on their statistical replicative ability.

Prediction

In order to check the prediction ability of the models, a value of root-mean-square-percentage (RMSP) error is computed for each model with the latest three-year licensing data.

RMSP = 1/3 x
$$
\sqrt{(y_{1991}-y_h_{1991})^2+(y_{1992}-y_h_{1992})^2 + (y_{1993}-y_h_{1993})^2}
$$
 x 100%

where

y 1991 is the ownership rate of car or motorcycle in 1991, and

 y^h ₁₉₉₁ is the estimated ownership rate of car or motorcycle in 1991.

As the empirical results shown in Table 6, most models give fairly good performance for shortterm prediction except model IV. However, a long-run model should not be judged only by its success or failure to estimate short-term changes.

Table 6 Short-term prediction performance—RMSP(%)

Figure 1 and Figure 2 show the growth patterns of car and motorcycle ownership for the five models. In general, different models give different growth patterns. For unconstrained models, model I and model II, only the models without socio-economic variables are shown in Figure 1 and Figure 2. It is clear that the car or motorcycle ownership in the traditional logistic curve, model I, will approach a very high level in the long-run. However, the car ownership in the generalized logistic curve, model II, will grow much slower than that in model I; model II gives a hump shape for the motorcycle ownership instead of the S shape in model I. In practice, the growth pattern of model II seems much acceptable to transportation people in Taiwan than that of model I.

Because s_{ij} is a function of UNADU, NI, P_c , and P_m , all the growth patterns of constrained diffusion models shown in Figure 1 and Figure 2 are based on the prediction of these socioeconomic variables. In this study, time-series analysis techniques are first used to generate the estimates for these variables. Then, the estimates of $s_{ij}(t)$ is used to compute the estimates of $y_i(t)$.

PREDICTION OF MULTI-CLASS VEHICLE OWNERSHIP LEE & SHIAW

The trend of car ownership in model V will approach a saturation level of around 350 cars per thousand people after the year of 2005, but that in all other models will increase in the long-run. Moreover, the trend of motorcycle ownership in model V is going to be stable at the level of around 300 vehicles per thousand people after the year of 2000, but that in all other constrained diffusion models will decline in the long-run. In practice, the growth pattern provided by model V seems acceptable for transportation people in Taiwan.

Figure 1 Long-term growth patterns for car ownership

Figure 2 Long-term growth patterns for motorcycle ownership

TOPIC 15

TRAVEL CHOICE AND DEMAND MODELLING

Comparing the growth trends of model I with model II, it is clear that the generalized logistic curve may not give a S shape of growth and may provide many possible growth patterns. Comparing the growth trends of model I with model III, it is clear that constrained logistic curve with a variable saturation level may also provide many possible long-run trends. Comparing the growth trends of model II with model IV, the saturation level used in model IV has little effect on the growth patterns.

CONCLUDING REMARKS

The ownership structure of motor vehicles in Taiwan is quite different from that in the developed countries. The paper reports a process for building and testing diffusion models on the basis of traditional logistic curves, including the specification and estimation of saturation levels. As reported in the paper, the newly developed models give promising results in an empirical study with Taiwan's licensing data of car and motorcycle. Hence, this study claims that the modelling approach used in the paper is at least a useful addition to the arsenal of techniques for researchers and practitioners who wish to analyze the long-term trends of multi-class motor vehicle ownership.

ACKNOWLEDGMENTS

The authors thank the financial support of NSC 81-0301-H006-503 from the National Science Council, R.O.C.

REFERENCES

Andrews, L.C. (1991) *Introduction to Differential Equations with Boundary Value Problems,* Harper Collins Publisher, N.Y.

Button, K., Pearman, A.D. and Fowkes, A.S. (1980) *Car Ownership Modelling and Forecasting,* Gower, Aldershot, UK.

Button, K., Ngoe, N. and Hine, *J.* (1993) Modelling vehicle ownership and use in low income countries, *Journal of Transport Economics and Policy* 27, 51-67.

Casetti, E. (1969) Why do diffusion process confirms to logistic trends, *Geographical Analysis* 1, 101-105.

Jansson, J.O. (1989) Car demand modelling and forecasting, *Journal of Transport Economics and Policy* 23, 125-140.

Khan, M.A. and Willumsen, L.G. (1986) Modelling car ownership and use in developing countries, *Traffic Engineering & Control* 27, 554-560.

Mogridge, M.J.H. (1967) The prediction of car ownership, *Journal of Transport Economics and Policy* 1, 52-74.

Mogridge, M.J.H. (1983) *The Car Market: A Study of the Statics and Dynamics of Supply-Demand Equilibrium,* Pion, London.

Mogridge, M.J.H. (1989) The prediction of car ownership and use revisited, *Journal of Transport Economics and Policy* 23, 55-74.

Rogers, E.M. (1983) *Diffusion of Innovations,* The Free Press, NY.

Sonis, M. (1986) A unified theory of innovation diffusion, dynamic choice of alternatives, ecological dynamics, and urban/regional growth and decline, *Ricerche Economiche* XL, 696-723.

Tanner, J.C. (1977) Car Ownership Trends and Forecasts, *Report LR799,* Transport and Road Research Laboratory, Crowthorne, UK.