Applied Economics, 1993, **25**, 1287–1292

Modelling the consumer's decision to replace durable goods: a hazard function approach

JENNIE E. RAYMOND, T. RANDOLPH BEARD and DANIEL M. GROPPER

Department of Economics, College of Business, Auburn University, Auburn, AL 36849, USA

This article analyses the consumer's durable good replacement decision using hazard models. In contrast to the typical limited dependent variable model often used in durable good demand studies, hazard models allow for much richer relationships between the ages of durable goods and the probabilities of their replacement. To illustrate the technique, a recursive system consisting of a regression equation and a hazard model is used to examine home heating system replacement decisions by residential customers of a major southeastern US electric utility. The results indicate that overall system replacement rates decline over time, and that the probability of replacement for specific households depends negatively on the age of the head of household and the availability of natural gas, and positively on system age and higher than expected household energy use.

I. INTRODUCTION

Statistical analysis of the demand for durable goods is complicated by several factors. First, by virtue of their extended effective lifetimes, the acquisition and replacement of durable goods involves an element of timing which does not arise in typical demand studies. As durable goods age, the consumer is faced with the issue of deciding when (and if) to replace them, and such actions are often undertaken prior to irreparable failure. Hence, the timing of the replacement decision is critical. Second, because of their longevity, the typical consumer replaces products such as central space heating systems, refrigerators and automobiles infrequently, leading to some data-driven difficulties in analysing durable goods purchases with many conventional statistical techniques. Finally, since the demand for many durables is, like the demand for energy, a derived demand, numerous factors which play an important role in durable good purchase behaviour, such as consumer tastes for 'comfort', are impossible to observe directly.

This paper demonstrates an approach for the empirical study of durable replacement activity that addresses all three of the complicating factors mentioned above. This approach, which is based on the estimation of hazard (duration) models for durable goods replacement, offers a different and theoretically accommodating way to analyse durable good replacement decisions without reliance on logit, probit or other inherently static specifications. The procedure is illustrated by analysing the replacement of home heating systems among customers of a large US electric utility. The estimated results are useful in several ways. First, information is obtained on the importance of appliance age, energy use characteristics and other factors in the durable good replacement decision. Second, the results provide a means to estimate replacement rates over a wide range of time periods. Further, while the application focuses on home heating systems, the technique could be applied to many energy consuming durable goods.

II. PREVIOUS RESEARCH

Economists have long noted the value of distinguishing between the demand for a stock of durable goods, the

0003-6846

purchase of a durable good, and demand for the services provided by durables.[1] Depending on the goal of the estimation, one or another of these features comes to the fore. As the primary interest of the present study is the decision to replace durable goods, the discussion will focus primarily on the purchase decision itself, and how this has been modelled previously.[2]

Because consumers typically buy only one unit at a time of most durable goods, analyses of consumer purchase decisions on the individual level have usually utilized some type of limited dependent variable (LDV) statistical technique, typically probit or logit models.[3] The consumer's purchase decision is assumed to depend on various factors that economic theory suggests might influence the consumer's choice. Using the three part distinction mentioned above, these factors might include relevant prices, consumers' incomes, consumer taste variables and, in the case of derived demand, factors influencing the value of the services which the durable good will give the consumer.

The consumer's choice to replace an existing unit may also be influenced by the potential energy savings obtainable with a replacement. Since repairing and maintaining an older unit is often a plausible alternative to acquiring a new unit, the age of the unit to be replaced is obviously relevant. Although appliance age can be included in a typical LDV-type model, such a formulation is inherently restrictive. Since LDV models are useful only to predict replacements in the 'next' time period, they are of limited usefulness as models of complex dynamic processes.

III. HAZARD MODEL ESTIMATION OF DURABLE GOODS REPLACEMENT

Hazard models are dynamic, non-linear statistical models which can be used to estimate the effects of observable characteristics on the length of time until a discrete event occurs. While many early hazard estimations focused on medical, biological or engineering problems, the first use of a hazard model in economic analysis was by Lancaster (1979), who studied unemployment duration. Amemiya (1985), Kiefer (1988) and Lancaster (1990) provide comprehensive surveys of this approach.

Formally, a hazard model assumes that the length of time, t, until an event occurs is a random variable with density $f(t)$ and cumulative distribution $F(t)$. Two auxiliary functions are derived from the density: the survivor function, $1-F(t)$, which shows the probability that the event has not occurred by some particular time, and the hazard rate. The hazard rate, $h(t)$, is the likelihood that an event occurs at a particular time given that it has not occurred previously: $h(t)=f(t)/[1-F(t)]$.

In practice, the hazard rate is specified initially and the density and survivor function are derived from it. The hazard rate may be hypothesized to depend on observable exogenous variables and on time itself. The dependence of the hazard rate on time implies that, as time passes and the event does not occur, the probability that the event occurs changes. The flexibility with which hazard models handle the impact of time on probability constitutes a major advantage over typical LDV formulations.

Here we hypothesize a hazard function of the following form:

$$h(t, x)=\exp(\gamma t) * \exp(\boldsymbol{\beta}'\mathbf{x}) \tag{1}$$

where $\mathbf{x}$ is a vector of exogeneous variables, t is time, and γ and $\boldsymbol{\beta}$ are parameters to be estimated. The sign of γ will indicate whether the probability of replacement increases or decreases over time.[4]

The expression for the hazard given in Equation 1 allows the derivation of the conditional distribution of time to failure, $f(t, x)$, by the relation: $h(t, x)=f(t,x)/[1-F(t, x)]$. We note that

$$f(t, x)=h(t, x) * [1-F(t, x)] \tag{2}$$

and that

$$1-F(t, x)=\exp\{-\int h(u, x)\,du\} \tag{3}$$

The likelihood function is composed of two types of terms, the first type corresponding to those observations for

[1] See e.g. Philips (1987).

[2] Acquisition patterns for durable goods have been examined by Kasulis *et al.* (1979) and by Dickson *et al.* (1983). Consumer durable replacement decisions and the timing of replacement have been studied by Bayus (1991). Bayus and Carlstrom (1990) have analysed ways in which consumer durables can be grouped in order to better model customer preferences and purchase decisions.

[3] With multiple purchases, static approaches include the use of Poisson regression techniques (see Paull, 1978). The use of discrete choice models was pioneered by Farrell (1954). Other pioneering papers include Cramer (1962), Wu (1965) and Cragg (1971). This disaggregated approach is to be distinguished from aggregate analyses of durable good purchases, such as the approach of Chow (1960). Deaton and Muellbauer (1988) offer an extensive overview of theory and practice in this area.

[4] It is possible to assume that the hazard rate is a non-monotonic function of time, for example

$$h(t, \mathbf{x})=\exp(\gamma_1 t+\gamma_2 t^2) * \exp(\boldsymbol{\beta}'\mathbf{x})$$

This specification was also estimated, but the results were not significantly different from those reported here. These results are available from the authors upon request.

which the hazard was observed to occur ('completed spells'), and the second representing consumers observed not to have experienced the hazard. By simple substitution, the likelihood function written in terms of the hazard rate is

$$L=\prod_{i=1}^{N_1} h(t_i, x_i) * \exp\left\{-\int h(u_i, x_i)\, du_i\right\} * \prod_{i=1}^{N_2} \exp\left\{-\int h(u_i, x_i)\, du_i\right\} \tag{4}$$

where N_1 and N_2 are the number of complete and incomplete spells, respectively.

Having broadly outlined the estimation strategy, some modifications of this approach are considered which are made necessary both by the economic nature of consumer behaviour and by some special features of the data set.

The application involves space heating system replacement by a randomly selected sample of households buying electricity from Alabama Power Company, a large regional US utility. These households provided extensive information on energy usage and household characteristics as part of Alabama Power Company's 1990 Residential Customer Survey Program. In addition, the households' energy usage and billing records were made available to us in connection with this survey.

All customers in our sample had some form of home heating system during the survey period. Hence, all consumer purchases of heating systems during the sample period represented replacements. In addition to replacements due to irreparable failures of existing systems, it seems plausible that such replacements arose from two broad motives: reliability and performance advantages represented by new units and potentially great energy savings from modernization. This second motive makes obvious a conceptual issue suggested by the derived nature of heating system demand: the value to a household of replacing an older unit depends on individual tastes for energy and the services (e.g. warmth or comfort) energy provides. Households which have strong preferences for appliance services would value improvements in efficiency more highly, and would therefore be more likely to replace an older unit with a new one.

The logic outlined above suggests how households' 'unobservable' tastes for energy services could be incorporated into the estimation. Since a households' energy usage depends on both their tastes for energy services and their stock of energy consuming durables, and since all sample households face identical electricity tariffs, similar geographic circumstances, and highly similar appliance prices, observed household energy usage rates, when adjusted for household appliance stocks, offer an attractive way to quantify differences in tastes for appliance services and, by implication, the values households place on improved efficiency.

To include a demand intensity measure into our estimation, the following approach is taken. Each consumer's average monthly usage of electricity (over an end-of-sample, one-year period) is modelled as a linear function of the household's stock of energy using durables and exogenous factors, Z

$$U=\boldsymbol{\theta} Z+e \tag{5}$$

where U is average usage in kwh/month, $\boldsymbol{\theta}$ is a vector of parameters to be estimated, and e is a normal disturbance which is interpreted as representing idiosyncratic tastes. Let $\hat{U}$ be the predicted value for U and let $e=U-\hat{U}$ be the estimated residual. Then the intensity variable, v, is defined as

$$v=(U-\hat{U})/\hat{U}=\hat{e}/\hat{U}$$

Thus, v provides a measure of the percentage deviation in customer usage unexplained by house size, the stock of durables owned and other observable factors. Computed for each consumer individually, these demand intensity measures can be used as explanatory variables, thereby controlling for unobservable individual tastes.

Additionally, the nature of our data set requires a modification to account for the large differential in the quantity of information on appliance ages between customers who have replaced their heating systems and those who have not. In particular, for those who did not replace their systems in the 3 years prior to the survey date, information was obtained on the age of their current system, but none on the age of their previous system which, for some individuals, had been replaced. For those who did replace their heating system in those 3 years, the age of the replaced heating system was noted. Since the sample of ages of systems at replacement represents a random sample of the ages of all heaters at replacement, the likelihood function was modified by weighting observations for those who replaced their systems by the inverse probability of replacement, and then normalizing all weights to sum to one.[5]

IV. MODEL ESTIMATION AND RESULTS

Our sample was taken from the Alabama Power Company (1990), Residential Customer Survey and Alabama Power Company billing records. Only those individuals who own their homes, and for whom the home was the primary residence, are included in the hazard model estimation (534 out of the original 702 observations).

To estimate v, the demand intensity variable, average kwh/month are regressed on a set of variables designed to capture a variety of household characteristics. These variables are defined in Table 1, which presents the estimation results for the demand intensity regression. These results are

[5] The probability of replacement within the past 3 years is 11%, so these observations have an unnormalized weight of about 9.

Table 1. *Household electrical energy use (average kwh/month) regression results*

Variable	Coefficient	t-value
Constant	256.16	0.77
Water heater age (years)	−6.47	−1.35
Space heater age (years)	7.83*	2.11
Number of electric heat pumps	20.41	0.23
Number of electric central systems	142.63*	1.82
Number of electric window AC units	22.48	0.42
Number of gas central systems	−77.19	−0.25
Average age electric heat pumps (years)	41.85*	7.97
Average age electric central system (years)	3.95	0.59
Average age electric window units (years)	0.59	0.12
Average age gas central system (years)	−8.48	−0.22
House square footage (1000s)	0.07	1.64
Single family home dummy	−278.17*	−2.30
Duplex dummy	282.56	1.15
Tri- or quadruplex dummy	349.90	1.43
5–9 family home dummy	−237.07	−0.81
10+ family home dummy	−192.23	−0.85
Urban/rural dummy (=1 if urban)	−11.60	−0.19
Number living in home	100.86*	4.79
Resident owns home dummy	62.35	0.62
Natural gas water heater dummy	−278.82	−0.90
Electric water heater dummy	92.87	0.30
Bottled gas water heater dummy	−280.49	−0.86
Number of type 1 appliances	163.47*	5.44
Number of type 2 applicances	33.27*	2.85
$N=702$ observations		
$R^2=0.3168$		
$F=13.098$		

*Indicates significance at the 5% level. The omitted housing type is mobile homes. Type 1 appliances are dishwashers, clothes washers and dryers. Type 2 appliances are televisions, stereos, refrigerators and VCRs. The dependent variable is average kwh usage per month over a 1-year period immediately prior to the survey.

relatively intuitive and straightforward. It is noted that square footage and the numbers and ages of appliances typically have the expected effects on usage: big houses with older appliances use more energy. Additional significant effects are represented by heating system ages, multiple heating systems, the number of residents in the household, the housing type and the numbers of various appliances owned.

In the specification of the hazard function, the exogenous variables in addition to a constant term are the demand intensity variable v, the age of the head of the household, household income, a dummy variable indicating that the customer lives in an urban area, square footage of the home, a dummy variable indicating that natural gas service is available and a dummy variable indicating that the customer has a poor credit rating.[6]

Table 2 presents the results of the hazard model estimation using the exponential specification with linear time-dependence as outlined in Section III. While these estimates are used later in some illustrative calculations, it is worthwhile to examine these results in some detail.

It is first noted that, unsurprisingly, the passage of time is highly significant in explaining replacement behaviour: the older the system, the greater the probability that it will be replaced. Further, the larger the value of the demand intensity variable, the more likely is home heating system replacement, a result consistent with the theoretical prediction. Two other factors are statistically significant indicators of replacement: the age of the head of the household, and the availability of natural gas service. The older the head of the household, the less likely he or she is to replace the heating system. Natural gas availability is also associated with a significantly lower probability of obtaining a new heating system, a result that may stem from possible differentials

[6] Households not reporting system replacement in the previous 3 years were treated as incomplete spells. Most of the variables included in the hazard function are largely self-explanatory. While the energy intensity variable has been described above, the poor credit rating dummy variable was taken directly from the utility company's rating system. All customers face the same electricity price schedule and, living in the same region, presumably face similar appliance prices except for urban/rural differentials.

Table 2. *Hazard model coefficients*

Variable	Coefficient	t-value
Constant	−0.29*	−1.77
v: unexplained energy usage (per 10 percentage points)	0.14*	1.89
Age of head of household (per 10 years)	−0.13**	−4.41
Income (per $10 000)	−0.02	−0.76
Urban/rural dummy (=1 if urban)	−0.16	−1.46
House square footage (per 1000)	−0.01	−1.51
Natural gas availability dummy	−0.34**	−3.31
Poor credit rating dummy	−0.22	−1.47
Time (years)	0.28**	7.26
$N = 534$ observations		

Note: Positive (negative) coefficients imply decreases (increases) in replacement time. t-values are asymptotic.
* Indicates significance at the 10% level.
** Indicates significance at the 1% level.

Table 3. *Marginal impacts on the probability of replacing the primary home heating system*

	Time period (years)			
Variable	1–3	4–6	7–9	1–20
v (change is 10 percentage points)	0.001	0.001	0.001	0.004
	(1.88)	(1.87)	(1.88)	(1.88)
Age of head of household (change is 10 years)	−0.009	−0.008	−0.007	−0.041
	(−3.81)	(−3.82)	(−3.83)	(−3.84)
Income (change is $10 000 in 1990 dollars)	−0.002	−0.001	−0.001	−0.007
	(−0.74)	(−0.75)	(−0.75)	(−0.75)
Urban	−0.011	−0.010	−0.008	−0.050
	(−1.38)	(−1.39)	(−1.39)	(−1.39)
Square feet (change is 1000 sq. ft)	−0.001	−0.001	−0.001	−0.004
	(−0.15)	(−0.15)	(−0.15)	(−0.15)
Natural gas available	−0.024	−0.020	−0.017	−0.104
	(−3.65)	(−3.65)	(−3.65)	(−3.65)
Poor credit rating	−0.016	−0.013	−0.011	−0.069
	(−1.41)	(−1.41)	(−1.41)	(−1.41)
Overall probability of replacement	0.073	0.068	0.063	0.402
	(12.33)	(13.18)	(14.19)	(14.95)

t-values are in parentheses. Marginal impacts are evaluated at sample mean values.

in the effective lifetimes of gas versus electric powered systems.[7]

The advantage of hazard estimation over LDV procedures can be illustrated by making the following simple calculations. First, consider the marginal impacts of the regression variables on the probabilities of system replacement, calculated as the change in probability of replacement (over a specified future interval) per unit change in the explanatory factor. Table 3 presents these results for various time intervals. For example, a 10 year increase in the age of the head of the household reduces the probability of replacement within 20 years by a little over 4%, a highly significant effect. Other significant effects include demand intensity, with higher intensity increasing replacement probabilities, and natural gas availability, which reduces the likelihood of replacement by about 10% within 20 years.

Table 4 offers a similar analysis focused on the expected lifetimes of home heating systems, and how these estimated times until replacement are affected by changes in exogenous variables. Overall, the mean time to system replacement is 25 years.

However, high energy usage intensity reduces the expected lifetime, with a 10% increase in unexplained usage shortening time to replacement by almost 2 years. Similar

[7] Replacement systems may be either gas or electric powered. However, those without gas service would always replace an electric powered system.

Table 4. *Marginal impacts on the expected time to replacement of the primary home heating system (years)*

Variable	
v (change is 10 percentage points)	−1.99
Age of head of household (change is 10 years)	1.86
Income (change is $10 000 in 1990 dollars)	0.31
Urban	2.30
Square feet (change is 1000)	0.18
Natural gas available	4.78
Poor credit rating	3.15
Overall time to replacement (years)	25.24

Marginal impacts are evaluated at sample mean values.

intuitively appealing effects are noted for the other explanatory variables.

V. SUMMARY AND CONCLUSIONS

This paper introduces the statistical technique of hazard model estimation to the study of durable goods replacement behaviour. The results obtained in the hazard model estimation are intuitively plausible and of statistical significance. While older systems are significantly more likely to be replaced, older heads of household are significantly less likely to obtain new systems. The availability of natural gas service also significantly reduces the probability of replacement. Further, those households with electricity consumption in excess of predicted levels are significantly more likely to replace their systems even when the age and type of their current system(s) is taken into account. This result conforms to the expectations of economic theory regarding durable goods replacement, and is important in explaining replacement behaviour.

While data limitations required us to limit our analysis to home heating systems, the methodology and techniques developed here are potentially applicable to a wide variety of durable good acquisition and replacement issues. While virtually all individuals in our sample had some form of home heating system, so that any purchase of a new system was a 'replacement', the initial (non-replacement) acquisition of other, 'non-saturated' appliances could also be studied with the hazard methodology. In this way, saturation issues, which are often of interest to energy analysts, can be directly evaluated with the hazard model approach. Hazard models also allow the researcher to overcome the inherently static approach of limited dependent variable modelling techniques. Our results show that the overall probability of replacement and the marginal impact of changes in the independent variables on the probability of replacement change over time. These findings would have been unobservable with traditional limited dependent variable techniques. Thus, it is concluded that the hazard model approach outlined and applied in this paper provides a useful and complementary alternative to the techniques traditionally used to model consumer behaviour in replacing durable goods.

ACKNOWLEDGEMENTS

Funding support from the Auburn Utilities Research Center is gratefully acknowledged. We would like to thank Jim Anderson, Andy Barnett, Hugh Granade, Gibson Lanier, John McDavid and an anonymous referee for helpful comments, and Bob Bolen, Dave Townley, Marty Wade and Bob Wenger for help with the data. The usual caveat applies.

REFERENCES

Amemiya, T. (1985) *Advanced Econometrics*, Harvard University Press, Cambridge, MA.

Bayus, B. L. (1991) The consumer durable replacement buyer, *Journal of Marketing*, **55**, 42–51.

Bayus, B. L. and Carlstrom, C. C. (1990) Grouping durable goods, *Applied Economics*, **22**, 759–73.

Chow, G. (1960) Statistical demand functions for automobiles and their use for forecasting, in *The Demand for Consumer Durables*, Harberger A. C. (ed.), pp. 149–78. University of Chicago Press, Chicago.

Cragg, J. G. (1971) Some statistical models for limited dependent variables with applications to the demand for durable goods, *Econometrica*, **39**, 829–44.

Cramer, J. S. (1962) *A Statistical Model of the Ownership of Major Consumer Durables*, Cambridge University Press, Cambridge, MA.

Deaton, A. and Muellbauer, J. (1985) *Economics and Consumer Behavior*, Cambridge University Press, Cambridge.

Dickson, P. R. Lusch R. E. and Wilkie, W. R. (1983) Consumer acquisition priorities for home appliances: a replication and extension, *Journal of Consumer Research*, **9**, 432–5.

Farrell, M. J. (1954) The demand for motor cars in the United States, *Journal of the Royal Statistical Society*, Series A (General) **117**, 171–201.

Hawdon, D. (1992) *Energy Demand: Evidence and Expectations*, Surrey University Press, London.

Kasulis, J. J., Lusch, R. F. and Stafford, E. F. (1979) Consumer acquisition patterns for durable goods, *Journal of Consumer Research*, **6**, 47–57.

Kiefer, N. (1988) Economic duration data and hazard functions, *Journal of Economic Literature*, **26**, 646–79.

Lancaster, T. (1979) Econometric methods for the duration of unemployment, *Econometrica*, **47**, 939–56.

Lancaster, T. (1990) *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge, MA.

Nickell, S. (1979) *The Investment Decision of Firms*, Cambridge University Press, Cambridge.

Paull, A. (1978) A generalized compound Poisson model for consumer purchase panel data analysis, *Journal of the American Statistical Association*, **73**, 760–73.

Phlips, L. (1987) *Applied Consumption Analysis*. Revised Edition, North-Holland, Amsterdam.

Westin, R. (1975) Empirical implications of infrequent purchase behavior in a stock-adjustment model, *American Economic Review*, **65**, 384–96.

Wu, J. (1965) An empirical analysis of household's durable goods expenditure, *Econometrica*, **33**, 761–80.