CONSUMER ANTICIPATIONS AND HOUSEHOLD CHARACTERISTICS

Thomas Stoterau, Bureau of the Census

I. Introduction

The intent of this paper is to examine the probability approach to ex-ante consumer behavior for different types of households. The general hypothesis to be explored is whether income or education, for example, affect the general suitability of purchase probability questions. In the past, any such concern has been parried by the assurance that the concept of chance is commonplace in a variety of American rituals. Horse racing, World Series' pools, Weather Bureau reports, and the "numbers" have always been cited as examples. It, admittedly, seems reasonable that nearly everyone should have encountered the idea of chance in at least one of these situations. Regardless, this argument provides no empirical evidence about the ability of all types of respondents to assess the likelihood of future personal events. Thus it is that we will look at auto purchase probabilties and subsequent purchase behavior in conjunction with income, education, type of family, and respondent.

In 1966 the Census Bureau's program to measure ex-ante purchase behavior pioneered a subjective probability methodology. The survey abandoned its intentions to buy questions for what seemed to be a rather logical extension of the intention's rationale. The new approach consisted of asking respondents to numerically describe their chances of purchasing various durables over several time horizons. This technique utilized an eleven point "flash card" (0,10,20,...,100). It was believed that this would allow respondents to rationally speculate about their purchase probabilities. Whereas responses had previously been restricted to yes-no (0,1), the likelihood of the purchase event could now be described with greater discrimination. This approach has continued as the central feature of the Survey of Consumer Buying Expectations ever since.

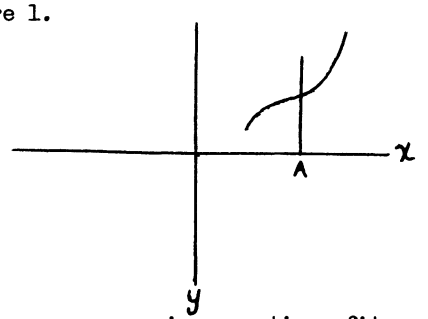
II. Description of Data and Analysis

The data used in this paper were gathered by the Census Bureau from a special panel known as the Consumer Anticipations Survey (CAS). The CAS study was developed as an experimental effort which would be used to revise and strengthen the Bureau's quarterly survey of consumer anticipations. Both the Census Bureau and the National Bureau of Economic Research participated in the design of the survey and questionnaires. In the study, about 3,500 households were asked detailed questions about their economic and demographic characteristics. These households were visited five times at roughly six-month intervals between May 1968 and November 1970. The sample selection was non-random and purposive. All of the selected census tracts were suburban and had moderately high to high median incomes. Such areas were purposely selected because these households were likely to yield a high frequency of positive responses to questions about economic activity.

The CAS survey provided information on a wide range of items concerning household decision making, and was the basis of an invited session at the 1971 ASA meetings in Fort Collins.1/

The analytical process which supports this paper is cross-sectional. That is, each household is considered an observation, and the dependent variable is the actual purchase of a new car (0,1) in a given six-month period. The independent variables, of course, vary with each equation. This is the standard, cross-section regression analysis. The work presented below differs from previous analytical efforts in that here the regression equations are run on partitioned independent variables.2/ In this case, the reasons for partitioning an independent variable are: (1) to examine the fit of the regression equation and (2) to observe the impact of the partitioned variable, over two or more bounded areas, upon the other variables in the equation. The first hopes that by segmenting a curvilinear relationship, we can find more nearly linear situations. Figure 1 presents a graphical illustration of a hypothetical case.

Figure 1.



As can be seen, regression equations fit over the two segments are more nearly appropriate than a single equation fit over the entire range of values. If, however, the relationship between the dependent and the independent variable is linear, the equations fit over the bounded segments will approximate the one fit over the entire x-axis. But even if a given partition does not improve the fit of the regression equation, it may still be useful analytically as a means of telling us about the impact of the partitioned variable upon the other variables. In this study, households were partitioned into those with incomes above $10,000 and those $10,000 and below. Regression equations with the same variables were then run for each of them. Other partitioned variables were treated similarly. All of the variables for these equations are defined in table 1, below.

III. Income and Anticipations

The selection of a partition is arbitrary. In this case, $10,000 was selected because it is about the median U.S. family income. The regression results for the $10,000 income partition can be seen in tables 2 and 3 at the end of the paper. In general, the R-squares are about the

same for both groups. The six-month and twelve-month purchase probabilties ($_6PP^*$ and $_{12}PP^*$) are much more significant for the upper income segment. That is, $_6PP^*$ and $_{12}PP^*$ explain much more of the variance in new car purchases for households with incomes above $10,000 than those under. The t-values are almost twice as large for the upper group.

Total family income ($I_{A,B}$) and other income ($OI_{A,B}$) are also more significant for the upper income segment. In contrast, the type of family ($TF_{A,B}$) is not significant for households above $10,000, but it is significant for the lower group. Since most of the households were husband - wife families, the type of family variable is almost a scaled variable for the number of children in the family (see Table 1). The respondent variable ($RS_{A,B}$) is also more significant for the lower group. Somewhat surprisingly, none of the other variables were significant for either group. These included the probability of the head being unemployed, education of head, and whether the household was currently making payments on a car. These results would suggest that the purchase probability method may be more suitable for households with incomes above $10,000. It does not, however, suggest that the approach is necessarily inappropriate for other households.

IV. Education and Anticipation

The first attempts to partition the education variable were not fruitful. These efforts first used four years of high school and then one to three years of college. Later, a partition at four or more years of college yielded some worthwhile results. The regression equations and their related coefficients can be seen in tables 4 and 5 at the end of this paper. One of the most interesting findings is that a dummy variable for respondent is not significant for households with a head with four or more years of college, but it is significant with a positive sign for those with less education. The respondent variable assumed a value of 1 if the head was involved in the interview and a zero if the head was not. This implies that for households with college educated heads it does not matter who the respondent is,but that for less well educated households it does. There may be implications in this finding which affect the on-going Survey of Consumer Buying Expectations (CBE). Our regular CBE survey accepts any person living in the household who is eighteen or older. It is generally believed that it is too difficult and expensive to restrict the respondent to one specified member of the households in a fairly large scale, current program. And despite the findings above reported, it is very unlikely that the benefit would warrant the cost of restricting the respondent.

As can be seen in table 4, below, the purchase probabilties ($_6PP^*$ and $_{12}PP^*$) are more significant for the upper education segment. In contrast, type of family ($TF_{A,B}$) and total family income ($I_{A,B}$) are generally more significant for the households whose head did not have four years of college. None of the other variables explained a significant portion of the variance in any of the equations.

Table 1.--CAS VARIABLES INCLUDED IN REGRESSION ANALYSIS

$RS_{A,B}$ - Respondent in interview (subscript refers to whether it is an above or below partition household)

1 = Head involved in interview
0 = otherwise

$EH_{A,B}$ - Education of head

1 = no education
2 = 1 to 8 years of elementary school
3 = 1 to 3 years of high school
4 = 4 years of high school
5 = 1 to 3 years of college
6 = 4 years of college
7 = 5 or more years of college

$U^*_{A,B}$ - Probability of head being unemployed in the next twelve months
(0, 10, 20,..., 100)

$OI_{A,B}$ - Other income for the household (interest and dividends, capital gains, rent, own business, and pensions)

Actual dollar amount (5 digits)

$I_{A,B}$ - Total annual income from all sources

Actual dollar amount (5 digits)

$TF_{A,B}$ - Type of family

1 = single male head, no children
2 = single male head, 1 or more children
3 = single female head, no children
4 = single female head, 1 or more children
5 = husband and wife, no children
6 = husband and wife, 1 child
7 = husband and wife, 2 children
8 = husband and wife, 3 children
9 = husband and wife, 4 or more children

$MP_{A,B}$ - Presently making payments on a car

0 = No
1 = Yes

$PG_{A,B}$ - Present grade in school of oldest child (if under 21)

00 to 18

$_6PP^*_{A,B}$ - Probability of purchasing a new car in the next six-months
(0, 10, 20,..., 100)

$_{12}PP^*_{A,B}$ - Probability of purchasing a new car in the next twelve-months

V. Conclusion

Contrary to my a priori suspicions, the regression equations for the various segments of the partitioned variable did not differ greatly. As is often the case with cross-section studies, all of the R-squares are disappointing. And due to the nature of the data, comparisons among the standard errors of the estimate (Sy.x) are not very meaningful. We did find that the purchase probabilties have larger t-values for households with incomes above $10,000 and for households whose head has four years of college. It also seemed possible that we would find that for some types of households, purchase probabilties are not significant. This did not happen. The t-values of $_6PP^*$ and $_{12}PP^*$ are fairly healthy in all of the equations.

Footnotes

1/ See Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C., pp. 126-170

Table 2.--REGRESSION EQUATIONS FOR HOUSEHOLDS WITH INCOME ABOVE $10,000 WITH OBJECTIVE AND ANTICIPATORY VARIABLES - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation number	R^2_A	REGRESSION COEFFICIENTS										$\bar{Y}$	$S_{y.x}$
		RS_A	EH_A	U^*_A	OI_A	I_A	TF_A	$_6PP^*_A$	$_{12}PP^*_A$	MP_A	PG_A		
I.....	.112	.012 (1.96)	-.008 (1.79)	-.004 (1.02)	-	-	-	.035 (18.98)	-	-	-	.104	.295
II....	.099	-	-.008 (1.75)	-	.0000 (3.03)	-	-	-	.0254 (17.05)	-.008 (.67)	.0052 (.389)	.087	.297
III...	.153	.004 (.746)	-002 (.39)	-034 (.79)	-	-	-	.044 (22.86)	-	-	-	.066	.295
IV....	.012	-	-.0035 (.704)	-	-	.0000 (6.00)	-.0004 (5.76)	-	-	-	-	.081	.319
V.....	.100	.0089 (1.49)	-	-.025 (.588)	-	.0000 (3.53)	.0006 (.090)	-	.0249 (16.62)	-	-	.007	.296

NOTE: The A subscript indicates that the variable is for households above the income partition.

Table 3.--REGRESSION EQUATIONS FOR HOUSEHOLDS WITH AN INCOME OF $10,000 OR LESS WITH OBJECTIVE AND ANTICIPATORY VARIABLES - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation number	R^2_B	REGRESSION COEFFICIENTS										$\bar{Y}$	$S_{y.x}$
		RS_B	EH_B	U^*_B	OI_B	I_B	TF_B	$_6PP^*_B$	$_{12}PP^*_B$	MP_B	PG_B		
I.....	.124	.0236 (2.19)	.0011 (1.42)	.0018 (.297)	-	-	-	.0439 (8.58)	-	-	-	.058	.255
II....	.126	-	-.003 (.381)	-	.000011 (.378)	-	-	-	.0331 (9.19)	-.025 (1.13)	.005 (.389)	.052	.255
III...	.049	.0097 (.934)	.0025 (.391)	-.0079 (1.38)	-	-	-	.0248 (5.17)	-	-	-	.036	.247
IV....	.063	-	-	-	-	-	-	-	.023 (6.41)	.0069 (3.24)	-	.0378	.245
V.....	.138	.0235 (2.21)	-	.0013 (.230)	-	.000009 (.764)	-.022 (2.13)	-	.0328 (9.24)	-	-	.065	.253

NOTE: The B subscript indicates that the variables is for households below the income partition.

Table 4.--REGRESSION EQUATIONS FOR HOUSEHOLDS WITH FOUR OR MORE YEARS OF COLLEGE - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation numbers	R^2_A	REGRESSION COEFFICIENTS										$\bar{Y}$	$S_{y.x}$
		RS_A	EH_A	U^*_A	OI_A	I_A	TF_A	$6^{PP^*}_A$	$12^{PP^*}_A$	MP_A	PG_A		
I.....	.133	.0056 (.77)	-	-.0039 (.62)	-	-	-	.0375 (16.71)	-	-	.0040 (.308)	.088	.284
II....	.111	.0042 (.57)	-	-.0042 (.659)	-	.000009 (1.604)	.0011 (1.37)	-	.0265 (14.47)	-	-	-.0134	.287
III...	.086	-	.0014 (.102)	-	-	-	.0081 (.997)	-	.0223 (13.13)	-	-	-.0151	.291
IV....	.0221	-	-.0134 (.83)	-	.000014 (1.51)	-	-	.0149 (6.00)	-	-	-	.077	.311
V.....	.058	.0054 (.72)	-	-.0034 (.53)	-	.000010 (2.33)	-	-	.0156 (9.75)	-	-	.037	.300
VI....	.0096	-	-	-	-	.000018 (3.875)	-.0090 (1.047)	-	-	.0199 (1.39)	-	-.042	.306

NOTE: The A subscript indicates that the variable is for households above the education partition.

Table 5.--REGRESSION EQUATIONS FOR HOUSEHOLDS WHOSE HEAD HAS LESS THAN FOUR YEARS OF COLLEGE - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation numbers	R^2_B	REGRESSION COEFFICIENTS										$\bar{Y}$	$S_{y.x}$
		RS_B	EH_B	U^*_B	OI_B	I_B	TF_B	$6^{PP^*}_B$	$12^{PP^*}_B$	MP_B	PG_B		
I.....	.096	.0226 (2.996)	-	.0002 (.512)	-	-	-	.0334 (12.77)	-	-	-.0185 (1.42)	.093	.293
II....	.104	.0194 (2.58)	-	-.0003 (.076)	-	.0000 (3.73)	-.0135 (1.88)	-	.0245 (11.82)	-	-	.037	.292
III...	.059	-	-.002 (.213)	-	-	-	-.0197 (2.67)	-	.0183 (10.02)	-	-	.105	.299
IV....	.007	-	.186 (1.15)	-	.0000 (1.34)	-	-	.008 (2.86)	-	-	-	.078	.311
V.....	.062	.021 (2.79)	-	-.0000 (.004)	-	.0000 (4.30)	-	-	.0136 (7.88)	-	-	.035	.299
VI....	.008	-	-	-	-	.0000 (3.56)	.0008 (.11)	-	-	.005 (5.64)	-	.032	.293

NOTE: The B subscript indicates that the variable is for households below the education partition.