ESTIMATION OF RELATIONSHIPS FOR LIMITED DEPENDENT VARIABLES[1]

BY JAMES TOBIN

"What do you mean, *less* than nothing?" replied Wilbur. "I don't think there is any such thing as *less* than nothing. Nothing is absolutely the limit of nothingness. It's the lowest you can go. It's the end of the line. How can something be less than nothing? If there were something that was less than nothing then nothing would not be nothing, it would be something—even though it's just a very little bit of something. But if nothing is *nothing*, then nothing has nothing that is less than *it* is."

E. B. White, *Charlotte's Web*
(New York: Harper, 1952) p. 28.

1. INTRODUCTION

IN ECONOMIC SURVEYS of households, many variables have the following characteristics: The variable has a lower, or upper, limit and takes on the limiting value for a substantial number of respondents. For the remaining respondents, the variable takes on a wide range of values above, or below, the limit.

The phenomenon is quite familiar to students of Engel curve relationships showing how household expenditures on various categories of goods vary with household income. For many categories—"luxuries"—zero expenditures are the rule at low income levels. A single straight line cannot, therefore, represent the Engel curve for both low and high incomes. If individual households were identical, except for income level, the Engel curve would be a broken line like OAB in Figure 1. But if the critical income level OA were not the same for all households, the average Engel curve for groups of households would look like the curve OB. A similar kind of effect occurs under rationing of a consumers' good. The ration is an upper limit; many consumers choose to take their full ration, but some prefer to buy less.[2]

[2] For a theoretical exposition of the effects on aggregate demand functions of lower or upper limits on individual expenditure in combination with differences in tastes among households, see [4] and [6] and the literature there cited.
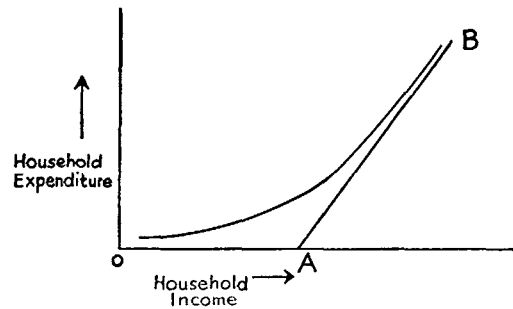
FIGURE 1

As a specific example, many—indeed, most—households would report zero expenditures on automobiles or major household durable goods during any given year. Among those households who made any such expenditure, there would be wide variability in amount.[3]

In other cases, the lower limit is not necessarily zero, nor is it the same for all households. Consider the net change in a household's holding of liquid assets during a year. This variable can be either positive or negative. But it cannot be smaller than the negative of the household's holdings of liquid assets at the beginning of the year; one cannot liquidate more assets than he owns.

Account should be taken of the concentration of observations at the limiting value when estimating statistically the relationship of a limited variable to other variables and in testing hypotheses about the relationship. An explanatory variable in such a relationship may be expected to influence both the probability of limit responses and the size of non-limit responses. If only the probability of limit and non-limit responses, without regard for the value of non-limit responses were to be explained, *probit analysis* provides a suitable statistical model. (See [2].) But it is inefficient to throw away information on the value of the dependent variable when it is available. If only the value of the variable were to be explained, if there were no concentration of observations at a limit, *multiple regression* would be an appropriate statistical technique. But when there is such concentration, the assumptions of the multiple regression model are not realized. According to that model, it should be possible to have values of the explanatory variables for which the expected value of the dependent variable is its limiting value; and from this expected value, as from other expected values, it should be possible to have negative as well as positive deviations.

A hybrid of probit analysis and multiple regression seems to be called for, and it is the purpose of this paper to present such a model.

2. THE MODEL

Let W be a limited dependent variable, with a lower limit of L. The limit may not be the same for all households in the population. Let Y be a linear combina-

[3] For figures on frequency of purchases and on the distribution of amounts spent among purchasers, see [1, Part II, Supplementary Tables 1, 5, and 10].

tion of the independent variables $(X_1, X_2, \cdots, X_m)$, to which W is by hypothesis related.

$$(1) \qquad Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_m X_m.$$

Households differ from each other in their behavior regarding W for reasons for which differences in the independent variables X and the lower limit L do not fully account. Those other differences are taken to be random and to be reflected in ϵ, a random variable with mean zero and standard deviation σ, distributed normally over the population of households. Household behavior is then assumed to be as follows:

$$(2) \qquad \begin{aligned} W &= L & (Y - \epsilon < L), \\ W &= Y - \epsilon & (Y - \epsilon \geqq L). \end{aligned}$$

Let $P(x)$ represent the value of the cumulative unit-normal distribution function at x; let $Q(x) = 1 - P(x)$; let $Z(x)$ be the value of the unit-normal probability density function at x. The distribution of $W - L$ may be derived from the distribution of ϵ, as follows:

For given values of the linear combination Y and the limit L,

$$(3) \qquad Pr(W = L \mid Y, L) = Pr(\epsilon > Y - L) = Q\{(Y - L)/\sigma\}.$$

$$(4) \quad Pr(W > x \geqq L \mid Y) = Pr(Y - \epsilon > x) = Pr(\epsilon < Y - x) = P\{(Y - x)/\sigma\}.$$

Consequently, the cumulative distribution function for W, for given Y and L, is:

$$(5) \qquad \begin{aligned} F(x; Y, L) &= 0 & (x < L), \\ F(L; Y, L) &= Q\{(Y - L)/\sigma\}, \\ F(x; Y, L) &= Q\{(Y - x)/\sigma\} & (x > L). \end{aligned}$$

The corresponding probability density function is:

$$(6) \qquad f(x; Y, L) = \frac{1}{\sigma} Z\{(Y - x)/\sigma\} \qquad (x > L).$$

The expected value of W for given values of Y and L is:

$$E(W; Y, L)$$

$$= LQ\{(Y - L)/\sigma\} + \int_L^\infty \frac{x}{\sigma} Z\{(Y - x)/\sigma\} \, dx$$

$$= LQ\{(Y - L)/\sigma\} + Y \int_{-\infty}^{(Y-L)/\sigma} Z(x) \, dx + \sigma \int_{-\infty}^{(Y-L)/\sigma} - xZ(x) \, dx.$$

Since $- xZ(x) = Z'(x) = dZ(x)/dx$, we have:

$$(7) \qquad E(W; Y, L) = LQ\{(Y - L)/\sigma\} + YP\{(Y - L)/\sigma\} + \sigma Z\{(Y - L)/\sigma\}.$$

3. THE MAXIMUM LIKELIHOOD SOLUTION

A sample includes q observations where W is at the limit L. Each observation consists of a limit L'_i, to which the dependent variable W'_i is equal, and a set of values of the independent variables $(X'_{1i}, X'_{2i}, \cdots, X'_{mi})$, where i is a subscript to denote the observation and runs from 1 to q. A sample also includes r observations for which W is above the limit L;[4] each one may be described as $(W_j, L_j, X_{1j}, X_{2j}, \cdots, X_{mj})$ where j runs from 1 to r.

Let $(a_0, a_1, a_2, \cdots, a_m, a)$ be estimates of $(\beta_0/\sigma, \beta_1/\sigma, \beta_2/\sigma, \cdots, \beta_m/\sigma, 1/\sigma)$. Let $I'_i = Y'_i a = a_0 + a_1 X'_{1i} + a_2 X'_{2i} + \cdots a_m X'_{mi}$ and let $I_j = Y_j a = a_0 + a_1 X_{1j} + a_2 X_{2j} + \cdots a_m X_{mj}$.

The likelihood of a sample is:

$$
\begin{aligned}
\phi(a_0, a_1, \cdots, a_m, a) &= \prod_{i=1}^{q} F(L'_i ; Y'_i, L'_i) \cdot \prod_{j=1}^{r} f(W_j ; Y_j, L_j) \\
&= \prod_{i=1}^{q} Q\left(\frac{Y'_i - L'_i}{1/a}\right) \cdot \prod_{j=1}^{r} aZ\left(\frac{Y_j - W_j}{1/a}\right) \\
&= \prod_{i=1}^{q} Q(I'_i - aW'_i) \cdot \prod_{j=1}^{r} aZ(I_j - aW_j).
\end{aligned}
$$
(8)

The natural logarithm of ϕ,

$$
\begin{aligned}
\ln \phi &= \phi^*(a_0 a_1, \cdots, a_m, a) \\
&= \sum_{i=1}^{q} \ln Q(I'_i - aW'_i) + r \ln a - \frac{r}{2} \ln 2\pi - \frac{1}{2} \sum_{j=1}^{r} (I_j - aW_j)^2.
\end{aligned}
$$
(9)

Let X_0 and X'_0 be identically 1 for all i and j. Then setting the derivatives of ϕ^* equal to zero gives the following system of $m + 2$ equations.

$$
\phi_k^* = \frac{\partial \phi^*}{\partial a_k} = \sum_{i=1}^{q} \frac{-Z(I'_i - aW'_i)X'_{ki}}{Q(I'_i - aW'_i)} - \sum_{j=1}^{r} (I_j - aW_j)X_{kj} = 0
$$
(10)
$$(k = 0, 1, 2, \cdots, m),$$

$$
\phi_{m+1}^* = \frac{\partial \phi^*}{\partial a} = \sum_{i=1}^{q} \frac{Z(I'_i - aW'_i)W'_i}{Q(I'_i - aW'_i)} + \frac{r}{a} + \sum_{j=1}^{r} (I_j - aW_j)W_j = 0.
$$

These equations are nonlinear. The quantity $-Z(x)/Q(x)$ is tabulated as $\Delta_{\min}$ in [2, pp. 185–88], where the argument for the table is $x + 5$.

The matrix of second derivatives, obtained by differentiating (10) is given by (11). Here $w'_{\min}(x)$ is the derivative of $-\Delta_{\min}(x)$, and may, like $\Delta_{\min}$, be found

---

[4] The value of L is assumed to be observable for the whole sample. I have made no investigation of the problems that would be presented by taking it as an unknown to be estimated.

by entering the tables of [2, pp. 185–188] with the argument $x + 5$.

$$\phi^*_{k\,t} = \frac{\partial^2 \phi^*}{\partial a_k \partial a_t} = -\sum_{i=1}^{q} X'_{k\,i} X'_{t\,i}\, w'_{\min}(I'_i - aW'_i) - \sum_{j=1}^{r} X_{kj}\, X_{tj}$$

$$(k, t, = 0, 1, \cdots, m),$$

$$(11)\quad \phi^*_{k,m+1} = \frac{\partial^2 \phi^*}{\partial a_k\, \partial a} = \sum_{i=1}^{q} W'_i X'_{k\,i}\, w'_{\min}(I'_i - aW'_i) + \sum_{j=1}^{r} X_{kj}\, W_j$$

$$(k = 0, 1, \cdots, m),$$

$$\phi^*_{m+1,m+1} = \frac{\partial^2 \phi^*}{\partial a^2} = -\sum_{i=1}^{q} W'^2_i w'_{\min}(I'_i - aW'_i) - \frac{r}{a^2} - \sum_{j=1}^{r} W_j^2\,.$$

Newton's method (see [3]) for iterative solution of (10), also known as the "method of scoring" (see [5]), may be applied as follows: Let

$$(a_0^{(0)}, a_1^{(0)}, \cdots, a_m^{(0)}, \cdots, a_{m+1}^{(0)})$$

be a trial solution, where, for notational convenience, a_{m+1} represents what has previously been written as simply a. (The choice of an initial trial solution will be discussed below.) New estimates

$$(a_0^{(0)} + \Delta a_0,\ a_1^{(0)} + \Delta a_1,\ \cdots,\ a_m^{(0)} + \Delta a_m,\ a_{m+1}^{(0)} + \Delta a_{m+1})$$

can be found by solving the set of $m + 1$ linear equations (12) for the Δa, where all the ϕ^*_k are assumed to be linear between the trial solution and the real solution.

$$\phi^*_k(a_0^{(0)} + \Delta a_0,\ \cdots,\ a_m^{(0)} + \Delta a_m,\ a_{m+1}^{(0)} + \Delta a_{m+1}) = \phi^*_k(a_0^{(0)}, a_1^{(0)}, \cdots, a_m^{(0)}, a_{m+1}^{(0)})$$

$$+ \sum_{t=0}^{m+1} \Delta a_t\, \phi^*_{k\,t}(a_0^{(0)}, a_1^{(0)}, \cdots, a_m^{(0)}, a_{m+1}^{(0)}) = 0 \qquad (k = 0, 1, 2, \cdots, m + 1).$$

$$(12)\quad \sum_{t=0}^{m+1} \Delta a_t\, \phi^*_{k\,t}(a_0^{(0)}, a_1^{(0)}, \cdots, a_m^{(0)}, a_{m+1}^{(0)}) = -\phi^*_k(a_0^{(0)}, a_1^{(0)}, \cdots, a_m^{(0)}, a_{m+1}^{(0)}).$$

The process may be repeated with the new estimates as provisional estimates until the Δa are negligible.

If the final estimates a_k are used to evaluate the matrix of second derivatives (11) at the point of maximum likelihood, the negative inverse of that matrix gives large-sample estimates of the variances and covariances of the estimates a_k around the corresponding population parameters.

4. TESTS OF HYPOTHESES

Hypotheses about the relationship of W to one or more of the independent variables X may be tested by the likelihood-ratio method. Consider for example, the hypothesis that $\beta_1 = \beta_2 = \cdots = \beta_m = 0$. This is the hypothesis that neither the probability nor the size of nonzero responses depends on the X's. According to the hypothesis, there remain only two parameters, β_0 and σ, to be estimated so

as to maximize (9), which now becomes:

$$\phi^*(a_0, 0, 0, \cdots, 0, a) = \sum_{i=1}^{q} \ln Q(a_0 - aW_i')$$

(13)

$$-\frac{r}{2} \ln 2\pi + r \ln a - \frac{1}{2} \sum_{j=1}^{r} (a_0 - aW_j)^2.$$

The maximizing values of a_0 and a may be found by solving equations (10) similarly simplified by putting all other a_k equal to zero. If (13) is evaluated with these solutions, then the logarithm of the likelihood ratio λ is the difference between (13) and the value of (9) when it is maximized without the constraint of the hypothesis. The statistic $-2 \ln \lambda$ is for large samples approximately distributed by chi-square with m degrees of freedom. In similar fashion other hypotheses about subsets of the β's may be tested.

5. INITIAL TRIAL ESTIMATES

The speed of convergence of iteration by Newton's method depends, of course, on the choice of the initial trial estimates. The following procedure for finding initial estimates relies on a linear approximation of the function $-Z(x)/Q(x)$ or, in other words, on a quadratic approximation of $\ln Q(x)$. This approximation converts the first $m + 1$ equations of (10) into linear equations in the a_k for given a. These equations may be solved to give the a_k as linear functions of a. When
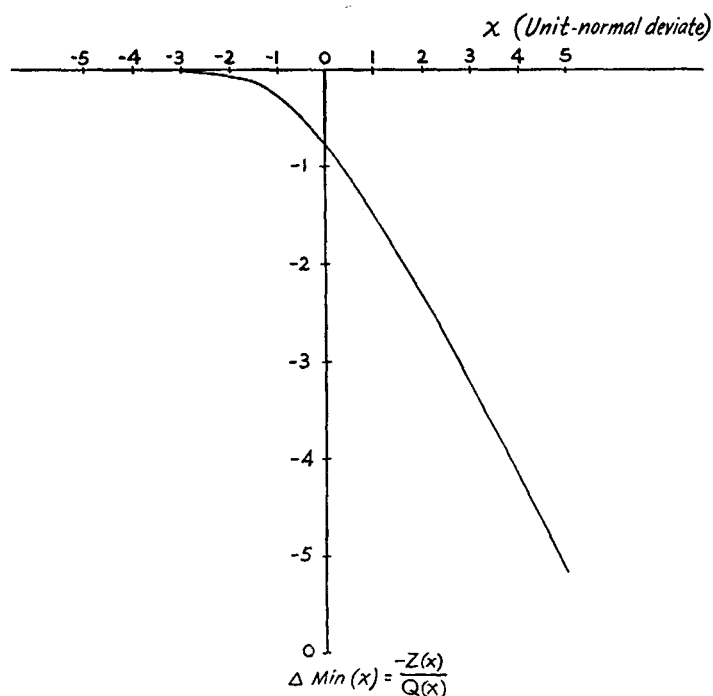

FIGURE 2

these solutions are substituted in the $m + 2$nd equation, it becomes a quadratic equation in a.

The function $\Delta_{\min}(x) = -Z(x)/Q(x)$ is graphed in Figure 2. Clearly it is not possible to approximate it linearly at all closely throughout its range. The important thing is to have a good approximation in the range where the sample observations are concentrated. There is no point in approximating well at $x = 4$ or $x = 5$ if the sample did not include constellations of the independent variables that made the probability of limit observations virtually nil. An easily computed estimate of the middle of the relevant range is x_o, the unit-normal deviate such that $Q(x_o) = q/(q + r)$, the proportion of cases in the sample for which the variable W takes on its limit-value. The tangent at x_o is one possible approximation, but not the best one, since it uniformly overestimates $\Delta_{\min}$ except at x_o itself.

Suppose that a linear approximation is fitted graphically:

$$(14) \qquad\qquad \Delta_{\min}(x) = A + Bx.$$

Substituting (14) in the first $m + 1$ equations of (10) gives

$$\sum_{i=1}^{q} (AX'_{ki} + Ba_o X'_{oi}X'_{ki} + Ba_1 X'_{1i}X'_{ki} + \cdots Ba_m X'_{mi}X'_{ki} - BaW'_i X'_{ki})$$

$$- \sum_{j=1}^{r} (a_o X_{oj}X_{kj} + a_1 X_{1j}X_{kj} + a_m X_{mj}X_{kj} - aW_j X_{kj}) = 0,$$

$$(15) \quad a_o \left[\sum_{j=1}^{r} X_{oj}X_{kj} - B \sum_{i=1}^{q} X'_{oi}X'_{ki} \right] + a_1 \left[\sum_{j=1}^{r} X_{1j}X_{kj} - B \sum_{i=1}^{q} X'_{1i}X'_{ki} \right]$$

$$+ \cdots + a_m \left[\sum_{j=1}^{r} X_{mj}X_{kj} - B \sum_{i=1}^{r} X'_{mi}X'_{ki} \right]$$

$$= a \left[\sum_{j=1}^{r} W_j X_{kj} - B \sum_{j=1}^{q} W'_i X'_{ki} \right] + A \sum_{i=1}^{q} X'_{ki} \qquad (k = 0, 1, 2, \cdots, m).$$

Solving (15) gives numbers g_k and h_k such that:

$$(16) \qquad\qquad a_k = g_k + h_k a \qquad\qquad (k = 0, 1, 2, \cdots, m).$$

The final equation of (10) is, after using the approximation of (14):

$$\frac{r}{a} + a_o \left[\sum_{j=1}^{r} X_{oj}W - B \sum_{i=1}^{q} X'_{oi}W'_i \right]$$

$$(17) \qquad + \cdots + a_m \left[\sum_{j=1}^{r} X_{mj}W_j - B \sum_{j=1}^{q} X'_{mi}W'_i \right]$$

$$- a \left[\sum_{j=1}^{r} W_j^2 - B \sum_{i=1}^{q} W'^2_i \right] - A \sum_{i=1}^{q} W'_i = 0.$$

When (16) is substituted in (17), it becomes a quadratic equation in a. The solution of (17) may then be used in (16) to obtain initial trial estimates of all the coefficients.

6. AN EXAMPLE

For purposes of illustration, an example has been worked out using data from the reinterview portion of the 1952 and 1953 Surveys of Consumer Finances conducted by the Survey Research Center of the University of Michigan for the Board of Governors of the Federal Reserve System.[5]

The data refer to 735 primary nonfarm spending units[6] who were interviewed twice, once in early 1952 and once in early 1953. The frequencies, averages, and other statistics for the reinterview sample should not be taken as representative of the population of the United States. The Surveys of Consumer Finances do collect data on distributions of income, liquid assets, and durable goods purchases that are representative of that population; tables on these distributions may be found in [1]. But the reinterview sample, on which the calculations of this paper are based, fails to be representative insofar as it omits spending units who moved between the two surveys. Moreover, these calculations are based on simple counts of sampled spending units, without allowance for the fact that the sampling design gave some spending units greater probabilities of being included in the sample than others. The purpose of this example is not to estimate population frequency distributions, but only to examine the relationship of durable goods expenditure to age and liquid asset holdings within this sample. It is not necessary to consider here how the relationship exhibited in this sample differs from the one that would be exhibited in a complete enumeration. But it may well be that the sample gives unbiased estimates of the parameters of the relationship, even though it gives biased estimates of the separate frequency distribution of the variables.

The variables are as follows:

W, the ratio of 1951–52 total durable goods expenditure to 1951–52 total disposable income. Durable goods expenditure is the two-year sum of outlays, net of trade-ins or sales, for cars and major household appliances and furniture. Two-year dis-

[5] A brief general description of the concepts and methods of the annual Surveys of Consumer Finances is given in Board of Governors of the Federal Reserve System, "Methods of the Survey of Consumer Finances," *Federal Reserve Bulletin*, July, 1950. For a more complete treatment, see also Klein, L. R., editor, *Contributions of Survey Methods to Economics*, New York: Columbia University Press, 1954. Reports of the 1952 and 1953 Surveys are given in Board of Governors of the Federal Reserve System, *1952 Survey of Consumer Finances*, reprinted with supplementary tables from *Federal Reserve Bulletin*, April, July, August, and September, 1952, and Board of Governors of the Federal Reserve System, *1953 Survey of Consumer Finances*, reprinted with supplementary tables from *Federal Reserve Bulletin*, March, June, July, August, and September, 1952.

[6] Of the 1036 spending units in the reinterview sample, these 735 have been the subject for calculations for other purposes and are therefore a convenient group to use in this analysis. Excluded are all spending units who had one or more of the following characteristics: (a) farm; (b) secondary, i.e., not the owner or principal tenant of the dwelling; (c) total income for the two years 1951–52 zero or negative; (d) not ascertained as to age of head of spending unit, amount of expenditure on durable goods during 1951–52, or amount of liquid asset holdings in early 1951. In addition, one extreme observation was excluded, where the spending unit has such a low positive two-year income that the ratio of durable goods expenditure and, especially, liquid asset holdings to income were very high.

posable income is the sum of the two annual incomes reported by the spending unit less estimated federal income tax liabilities. Both expenditure and income were reported for 1951 in the interview in early 1952, and for 1952 in the second interview, in early 1953. Since expenditure is necessarily zero or positive, and since zero and negative incomes have been excluded, the ratio is necessarily zero or positive.

X_1, *the age of the head of the spending unit*, as reported in 1953, on the following scale:

18–24 yrs.	1
25–34 yrs.	2
35–44 yrs.	3
45–54 yrs.	4
55–64 yrs.	5
65 or more years	6

TABLE I

SUMS OF SQUARES AND CROSS PRODUCTS

	183 limit observations				552 non-limit observations			
	$X_0' = 1$	X_1'	X_2'	W'	$X_0 = 1$	X_1	X_2	W
$X_0' \equiv 1$	183				$X_0 \equiv 1$ 552			
X_1'	824	4056			X_1 1976	8060		
X_2'	102.15	552.03	402.3333		X_2 168.06	751.54	255.6740	
W'	0	0	0	0	W 61.449	207.598	20.559	13.113087

TABLE II

ITERATIVE ESTIMATION OF PARAMETERS

	a_0	a_1	a_2	a
Initial trial values	1.326	−.2200	.0330	7.984
First derivatives	−4.398	−21.759	−1.812	.898
Second derivatives a_0	−680.557			
a_1	−2542.416	−10,805.486		
a_2	−238.41	−1128.653	−535.223	
a	61.449	207.598	20.559	−21.772
Indicated changes	.0152	−.00507	.00199	.0376
Second trial values	1.3407	−.2251	.0350	8.022
First derivatives	−.047	.292	.064	.002
Second derivatives a_0	−680.260			
a_1	−2540.666	−10,795.522		
a_2	−238.26	−1,127.900	−535.893	
a	61.449	207.598	20.559	−21.688
Indicated changes	−.0015	.00037	.00001	−.00064
Final estimates	1.3392	−.2247	.0350	8.022
Standard errors	(.118)	(.0295)	(.0495)	(.252)

TABLE III

ITERATIVE ESTIMATION OF PARAMETERS ASSUMING THAT $\beta_2 = 0$

	a_0	a_1	a
Initial trial values	1.337	−.219	8.040
First derivatives	−2.841	−16.124	−.001
Second derivatives: a_0	−680.419		
a_1	−2541.428	10,799.12	
a	61.449	207.598	−21.652
Indicated changes	.010	−.004	−.010
Second trial values	1.347	−.223	8.030
First derivatives	−.456	−2.017	+.116
Second derivatives: a_0	−680.179		
a_1	−2539.988	−10,790.472	
a	61.449	207.598	−21.674
Indicated changes	.001	.0003	−.005
Final estimates	1.347	−.223	8.030
Standard errors	(.117)	(.028)	(.252)

X_2, *the ratio of liquid asset holdings at the beginning of 1951 to 1951–52 total disposable income.* Liquid asset holdings include bank deposits, savings and loan association shares, postal savings, and government saving bonds.

In this example, the lower limit L is zero for all cases. Table I shows the basic data.

Table II presents the estimates of the parameters obtained by the initial approximation and reports the successive iterations leading to the maximum likelihood estimates. Estimates are shown also in Table III, on the assumption that there is no relation between W and liquid asset holdings X_2.

In the approximation used to obtain initial trial values, the function $-Z(x)/Q(x)$ was approximated by the tangent at the point $x_o = .67$, so that $Q(x_o) = .25$,

TABLE IV

ESTIMATED VARIANCES AND COVARIANCES OF PARAMETER ESTIMATES

	a_0	a_1	a_2	a
a_0	+.0139			
a_1	−.00318	+.000867		
a_2	+.000880	−.000454	+.00245	
a	+.00987	−.00115	+.000470	+.0635

On assumption that $\beta_2 = 0$:

	a_0	a_1	a	
a_0	+.0136			
a_1	−.00302	+.000784		
a	+.00970	−.00106	+.0635	

TABLE V

CALCULATION OF EXPECTED VALUES

		$X_2 = 0$				$X_2 = 2$		
X_1	$I = 1.3392$ $- .2247X_1$	Calculated probability of buying $P(I)$	$Z(I)$	Calculated expected value $E(W) = (IP + Z)/8.022$	$I = 1.4092$ $- .2247X_1$	Calculated probability of buying $P(I)$	$Z(I)$	Calculated expected value $E(W) = (IP + Z)/8.022$
0	1.3392	.910	.163	.172	1.4092	.921	.148	.180
1	1.1145	.867	.214	.147	1.1845	.882	.197	.155
2	.8898	.813	.267	.123	.9598	.832	.252	.131
3	.6651	.747	.319	.102	.7351	.768	.304	.108
4	.4404	.670	.362	.082	.5104	.695	.350	.088
5	.2157	.585	.390	.064	.2857	.612	.383	.070
6	− .0090	.497	.399	.049	.0610	.524	.398	.054
7	− .2337	.408	.388	.037	− .1637	.435	.393	.040
8	− .4584	.323	.359	.026	− .3884	.350	.370	.029

the proportion of nonzero cases in the sample. Thus the constants A and B in (15) and (17) were equal to − .76003 and − .75771 respectively.

Estimates of the variances and covariances of the parameter estimates can be obtained from the negative of the inverse of the final matrix of second derivatives. These are shown in Table IV. The corresponding standard errors of the coefficients are given in the final rows of Tables II and III.

The size of the standard error of a_2 indicates that the hypothesis that $\beta_2 = 0$, that there is no net relationship between expenditure and liquid asset holding, cannot be rejected. This hypothesis can also be tested, with the same conclusion, by the likelihood-ratio method. At the point of maximum likelihood, unrestricted by this hypothesis, ϕ^* in (9) has the value $722.5 - (552/2) \ln 2\pi$. The final estimates in Table III correspond to the point of maximum likelihood restricted by the hypothesis that $\beta_2 = 0$. At this point ϕ^* has the value $721.8 - (552/2) \ln 2\pi$. The statistic $- 2 \ln \lambda$ is thus equal to 1.4, which is not a significant value of chi-square with one degree of freedom.
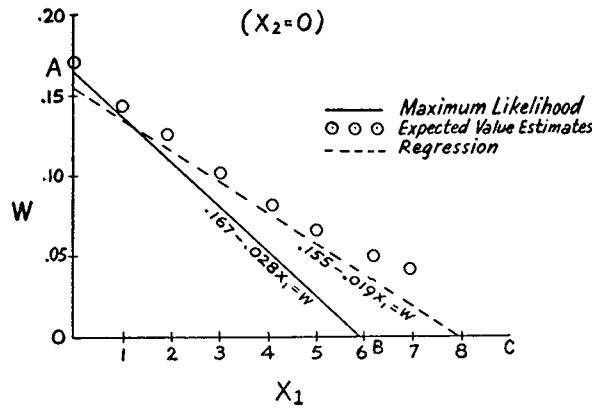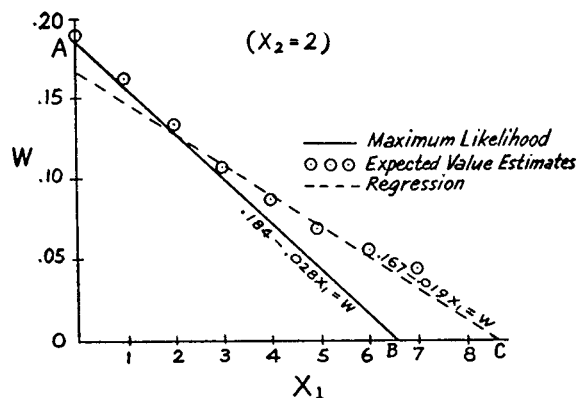

FIGURE 3a

FIGURE 3b

A test of the hypothesis that neither age nor liquid asset holding has any effect on expenditure on durable goods may also be made by the likelihood-ratio method. Assume, in accordance with the hypothesis that $\beta_1 = \beta_2 = 0$, that the values of a_o and a that maximize (13) are found to be .4839 and 7.720. For these values, $\phi^* + 552/2 \ln 2\pi$ is equal to 692.7. Hence $-2 \ln \lambda$ is equal to 59.6, a significant chi-square for two degrees of freedom. The hypothesis must be rejected. Thus this test, as well as the size of the estimated standard error of a_1, indicates a significant relationship of durable goods expenditure to age.

The relationship of W to X_1 and X_2, as estimated in Table II, is shown in Figure 3, as the broken line ABC. The expected value of W implied by this relationship may be computed from (7) in the manner illustrated in Table V. These points are also shown in Figure 3. For comparison, the least squares multiple regression of W on X_1 and X_2 has also been plotted. The estimated effect of liquid asset holding X_2 has been illustrated by drawing two graphs relating W to X_1, the first (Figure 3-a) on the assumption that $X_2 = 0$ and the second (Figure 3-b) on the assumption that $X_2 = 2$. Observed proportions buying and average values of W at various levels of X_1, without regard to X_2, are shown in Table VI.

The expected value locus, estimated by the method of this paper, is nonlinear.

TABLE VI

OBSERVED VALUES

(all values of X_2)

X_1	Proportion buying	Observed average value of W
1	.812	.124
2	.884	.118
3	.862	.098
4	.751	.083
5	.664	.145
6	.516	.047

It is always above the broken line ABC, asymptotic to AB at the left where the probability of not buying ($W = 0$) approaches zero, and asymptotic to BC at the right where the probability of buying ($W > 0$) approaches zero. Multiple regression approximates this nonlinear locus with a linear relationship. As Figure 3 shows, the approximation is fairly close for the central range of values of the sample. But outside the central range there can be large discrepancies. There are indeed conceivable values of the independent variables for which multiple regression would give negative estimates of the expected value of W. It is true that the absence of negative observations in the sample tends to keep the regression above the axis until extreme values of the independent variables are reached. But this protection is purchased at the cost of making the regression line so flat that expenditure is underestimated at the opposite end. These discrepancies could be important in predicting expenditure for extreme cases or for aggregates which include extreme cases.

Yale University

REFERENCES

[1] BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM: *1953 Survey of Consumer Finances*, reprinted with supplementary tables from *Federal Reserve Bulletin*, March, June, July, August, and September, 1953.

[2] CORNFIELD, J., AND N. MANTEL: "Some New Aspects of the Application of Maximum Likelihood to the Calculation of the Dosage Response Curve," *Journal of the American Statistical Association*, 45 (1950), 181–210.

[3] CROCKETT, J. B., AND H. CHERNOFF: "Gradient Methods of Maximization," *Pacific Journal of Mathematics*, 5 (1955), 33–50. (Reprinted as Cowles Commission New Series Paper No. 92.)

[4] FARRELL, M. J.: "Some Aggregation Problems in Demand Analysis," *Review of Economic Studies*, 21 (1954), 193–204.

[5] RAO, C. R.: *Advanced Statistical Methods in Biometric Research* (New York: John Wiley & Sons, 1952), 165–172.

[6] TOBIN, J.: "A Survey of the Theory of Rationing," *Econometrica*, 20 (1952), 521–553.