ESTIMATING PATTERNS OF SAVINGS BEHAVIOR FROM SAMPLE SURVEY DATA

BY LAWRENCE R. KLEIN

This paper attempts to estimate savings equations from sample survey data combining economic, demographic, and attitudinal variables. Savings as a function of income, family size, liquid assets, income change, and age of family head is estimated separately for home-owners and renters.

The residuals from these equations are then studied in relation to the perceived degree of permanence of income change, income expectations, and general economic outlook. In addition, equations are estimated for a pooled group of home-owners and renters, with a dummy home-ownership variable, and for different income change groups.

Cross-section data obtained in sample surveys provide a relatively unused source of information from which to construct empirical equations of household behavior. The fruitfulness of this source is that it enables us to build patterns of behavior from the ultimate microeconomic units who make basic decisions. In addition to their contributions to microeconomic analysis, these data enable us to study distributional effects (e.g., the effect of income or liquid asset distribution on savings); to use large-sample statistical theory; to study the frequency distribution of estimated random disturbances (residuals); to incorporate demographic, attitudinal, and other noneconomic variables in the behavior equations. To some extent, all these points represent limitations which cannot easily be overcome in empirical work with aggregative time series samples. In view of the continuing discussion of the appropriate structure of the "savings function," we turn to data gathered in the Surveys of Consumer Finances conducted by the Survey Research Center, University of Michigan, for the Board of Governors of the Federal Reserve System for new information that may throw light on the properties of this important functional relationship. A major objective of this paper is to explore the possibility of introducing sociological and psychological variables, together with the more traditional economic variables, in equations of household behavior. The main sociological variables used are demographic measures like family size, age of family head, home-ownership status, race, etc. By psychological variables we mean attitudinal expressions such as income expectations, general economic outlook, and price expectations. The traditional economic variables consist of savings, income, income change, liquid asset holdings, and debt.

The particular form of the savings equation used in this paper was chosen because it is a simple form consistent with the basic assumption of homoscedastic disturbances and reflects some of the important

1 Research under a grant from the Carnegie Corporation. This paper will be reprinted as Cowles Commission Paper, New Series, No. 55.
nonlinearities in observed behavior. A simple linear relation among the
variables listed above is unquestionably inadequate as a representation
of individual behavior in our sample. Our procedure has been to express
savings, a variable about which households make free decisions, in terms
of predetermined variables about which households do not, in general,
make free short-run decisions. The predetermined variables include
disposable income, beginning-of-period liquid assets, income change, age
of spending-unit head, and spending-unit size. Essentially, we assume
that the joint distribution of the predetermined variables is independent
of the parameters of the savings equation. We shall hold that this
assumption is reasonable and consequently that least-squares regressions
of a single endogenous variable on the predetermined variables provide
desirable estimates of structural parameters of the savings function.

The calculations presented here are all obtained from a special sample
of urban spending units who did not move between January–February,
1948, and January–February, 1949, and who were interviewed on the
two successive dates. This sample is known as the reinterview sample\footnote{The reinterview study of the Survey Research Center was carried out under
a grant from the Rockefeller Foundation.} and consists of 655 cases chosen so as to give a representative sample of
the urban areas of the United States. The reinterview sample is used by
the author for pilot study purposes and because it avoids memory errors
in obtaining responses to such variables as income change and liquid
asset holdings a year ago. Further work with independent one-time
samples of 3,000–3,500 cases is currently under way.

The sample of 655 is first divided into three groups: one group consists
of respondents from whom answers to some basic questions were not
satisfactorily obtained (called “N.A.’s”—not ascertained); a second
group consists of home-owners; and a third group consists of non-
home-owners—roughly the same as renters. The N.A. cases are not
analyzed. The distinction between renters and home-owners is made for
several reasons. In the first place, these two groups do constitute, in a
sense, different socioeconomic classes, although this is a less compelling
reason than the three to follow. Secondly, the initial conditions for the
two groups are different—one has housing wealth and the other does not.
Thirdly, mortgage holders among the home-owners have contractual
savings obligations in the form of debt repayment. Finally, the savings
and income concepts for home-owners do not account for depreciation
of homes and imputed rental income.

We begin our analysis with the following variables:\footnote{The reader is referred to articles on the Surveys of Consumer Finances in
the Federal Reserve Bulletin, 1948–1950, for precise definitions of all variables and
for details on the survey methods used.} $S$ represents
savings during 1948; $Y$, disposable money income during 1948; $N$, size
of the spending unit at the beginning of 1949; $L_{-1}$, liquid assets at the
beginning of 1948; and $a$, age of spending unit head at the beginning of
1949.

Estimated relationships are:

$$
S/Y = -0.93 + 0.35 \log Y/N - 0.21 \frac{L_{-1}}{Y}
$$

(1)

$$
+ 0.03 \frac{Y - Y_{-1}}{Y_{-1}} + 0.0013a + u_1,
$$

(0.08) \hspace{1cm} (0.02)

$$
R = 0.57, \quad S_u = 0.42;
$$

Renters (318 cases)

$$
S/Y = -1.61 + 0.48 \log Y/N - 0.25 \frac{L_{-1}}{Y}
$$

(2)

$$
+ 0.07 \frac{Y - Y_{-1}}{Y_{-1}} + 0.0054a + u_2,
$$

(0.11) \hspace{1cm} (0.03)

$$
R = 0.49, \quad S_u = 0.56.
$$

The numbers in parentheses are standard errors, $R$ is the multiple cor-
relation coefficient, and $S_u$ is the standard error of estimate. The main
distinction between the two groups is the fact that equation (1) shows a
higher constant term than (2), a difference that should be accounted for
largely by contractual savings of mortgage holders and the failure to
include house depreciation in saving estimates (negative savings) or
imputed rents in income estimates. These three items work in the
direction of causing the constant term to be higher in (1) than in (2).

All 606 observations can be pooled into a single group and analyzed
with respect to the same variables plus a dummy variable: $H = 1$ if
spending unit is a home-owner, $H = 0$ if spending unit is a renter.
The pooled sample yields the following estimate:

$$
S/Y = -1.36 + 0.41 \log Y/N - 0.23 \frac{L_{-1}}{Y}
$$

(3)

$$
+ 0.06 \frac{Y - Y_{-1}}{Y_{-1}} + 0.0038a + 0.13H + u_3,
$$

(0.07) \hspace{1cm} (0.02)

$$
R = 0.52, \quad S_u = 0.50.
$$

In all computations common logarithms are used. It should also be remarked,
for the benefit of those readers who attended the oral presentation of this paper
at the Chicago meeting of the Econometric Society in December, 1950, that
corrected estimates of sampling errors of parameters are given here. The original
estimates overstated the sampling errors. The regression equations are all weighted
regressions, with weights inversely proportional to net effective sampling rates.
These samples have been checked against the following three possible criticisms: (1) The group of N.A.'s, for which any one of the variables used is missing, may behave differently from the cases in the sample. We used estimated values of the missing variables assigned by the Survey Research Center on an independent basis and found that results are not significantly different from those obtained with the expanded sample. (2) Farmers, who present accounting difficulties in household surveys, have been automatically excluded from the reinterview sample, but urban business-owners, who present similar accounting difficulties, have not. In the renter group we calculated equations with and without business-owners and found the results not significantly different in the two cases. (3) Debt as well as assets constitutes part of the initial conditions—wealth holdings—and our equations neglect this fact. For the renter group we have introduced a debt variable and found that the other coefficients remain unchanged, while the coefficient of debt is near zero and is statistically unreliable. We do not rule out the possibility that debt is an important variable; the present sample is uninformative. Technical difficulties prevent us from getting beginning-of-year debt estimates for the home-owner group.

The estimated coefficient of income change is statistically unreliable, as is true of the age coefficient in the home-owner group. Mortgage payments tend to decrease with age; this could be partially responsible for the smaller and less reliable coefficient in equation (1). The fact that similar results show up in the renter and home-owner groups when independently treated does, however, lend some support to the estimates. One is also struck by the similarity of the coefficients of log \( Y/N \) and \( L_{-1}/Y \) in the two cases. The equations are still incomplete in the sense that there are further initial wealth conditions that have not been measured and used in the equations. The missing variables comprise stocks of durable consumer goods, currency, real estate, business net worth, stock shares, and nongovernment bonds. The renter group is relatively free from these deficiencies, especially for the calculations in which business owners are excluded; thus we do have a measure of confidence in our statistical findings even though the analysis is not yet airtight.

It appears in equations (1)–(3) that there is not much of an influence of income change on savings decisions. In one sense, we cannot reject the hypothesis that the coefficient of income change is zero, but the coefficient could also be as large as 0.15, or it could even be negative.

---

5 The same conclusion follows if log \( Y/N \) — log \( (Y/N)_{-1} \) is used instead of \( (Y - Y_{-1})/Y_{-1} \).

6 In the Surveys of Consumer Finances, liquid assets consist of personal bank accounts, U.S. Government Bonds, postal savings accounts, shares in savings and loan associations, and shares in credit unions.
According to one a priori hypothesis, consumers are slow in adjusting their levels of living to changed situations and tend to have high serial correlation in their expenditures. This behavior pattern would establish a positive correlation between income change and savings. The opposite hypothesis, that people spend income increases, particularly on durables, as they are received and contract expenditures in the face of income declines, cannot, however, be ruled out on any theoretical grounds. A very different sort of hypothesis is that consumers behave asymmetrically, reacting one way to income increases and another way to decreases. Equations (1)–(3) suggest that there are a variety of reactions when income changes, with some households varying positively and some negatively. This denies neither of the first two hypotheses. The third hypothesis requires that we reorganize our data. At the risk of obtaining too few observations, we have broken our sample into two more groupings—those spending units experiencing income decreases and those experiencing income increases. We shall not use the age variable in the calculations with these smaller groups.

The new statistical estimates are:

**Home-owners with income decreases (89 cases)**

\[
S/Y = -1.15 + 0.47 \log Y/N - 0.31 \frac{L_{-1}/Y}{Y_{-1}} - 0.54 \frac{Y - Y_{-1}}{Y_{-1}} + v_1, \tag{4}
\]

\[
R = 0.78, \quad S_* = 0.50;
\]

**Renters with income decreases (111 cases)**

\[
S/Y = -0.93 + 0.45 \log Y/N - 0.25 \frac{L_{-1}/Y}{Y_{-1}} + 1.61 \frac{Y - Y_{-1}}{Y_{-1}} + v_2, \tag{5}
\]

\[
R = 0.59, \quad S_* = 0.83;
\]

**Home-owners with income increases (199 cases)**

\[
S/Y = -0.68 + 0.26 \log Y/N - 0.06 \frac{L_{-1}/Y}{Y_{-1}} + 0.02 \frac{Y - Y_{-1}}{Y_{-1}} + v_2, \tag{6}
\]

\[
R = 0.26, \quad S_* = 0.34;
\]
Renters with income increases (207 cases)
\[
S/Y = -0.31 + 0.11 \log Y/N + 0.00003 L_{-1}/Y \\
(0.05) \hspace{1cm} (0.022)
\]
\[
+ 0.01 \frac{Y - Y_{-1}}{Y_{-1}} + u_i, \\
(0.02)
\]
\[R = 0.16, \quad S_r = 0.20.\]

Equations (4)–(7) are admittedly based on very small samples with the consequent magnification of standard errors of estimates; yet one finding remains unmistakably clear, namely, that income decrease intensifies the marginal effect of liquid asset holdings on saving. If a household experiences a drop in income and, at the same time, has liquid asset holdings, it is very likely to save less than households with the same income drop but smaller liquid assets or than households with the same liquid assets but no income drop. Evidence of an extremely important interaction effect is found here. This may prove to be of some significance in business cycle analysis. The author’s colleague, J. N. Morgan, suggests the likely hypothesis that the combination of many influences leads to extreme behavior. For example, if a household has simultaneously an income decline, substantial liquid assets, large hospital bills, extraordinary moving expenses, etc., it is almost certain to dissave large amounts or save much less than it would in the absence of these combined events.

The above findings have been reported elsewhere.\textsuperscript{7} In this paper it is intended to stress new results that push the analysis deeper and show some potentialities of the survey approach. Asymmetry in reaction to income change may not be wholly contained in different reactions to increases or decreases of income. The nature of the changes, especially the nature as perceived by the household experiencing the change, may lead to asymmetrical patterns of behavior. In trying to get a measure of perception of income change we are immediately led into the psychological problems of attitudinal analysis.

Following a technique introduced by G. Katona\textsuperscript{8} we shall group respondents simultaneously according to the two variables, income change between 1947 and 1948 and expected income change between 1948 and 1949. The expectations are determined from surveys conducted during January and February, 1949. The income change data are those used in


equations (1)–(7). Three classes of income change are used in the calculations to follow. They are: (1) those spending units whose disposable income increased by more than $50.00 between 1947 and 1948, denoted by +; (2) those units whose disposable income changed by $50.00 or less, denoted by 0; and (3) those units whose disposable income decreased by more than $50.00, denoted by −.

The four groupings for income expectations are: (1) those units who expected their income to rise from 1948 to 1949, denoted by +; (2) those units who expected their income to remain the same, denoted by 0; (3) those units who expected their income to decrease, denoted by −; and (4) those units whose income expectation is unknown, denoted by ?.

Writing income change for 1947–1948 first, and expected income change for 1948–1949 second, we group the respondents into the categories given in the table below.

<table>
<thead>
<tr>
<th>Category</th>
<th>(Y - Y_{-1})</th>
<th>Expected Change</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
<td>+</td>
<td>Permanent increase</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>−</td>
<td>−</td>
<td>Permanent decrease</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>−</td>
<td>Nonpermanent increase</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>−</td>
<td>+</td>
<td>Nonpermanent decrease</td>
</tr>
<tr>
<td></td>
<td>−</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Permanently unchanged</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>−</td>
<td>Nonpermanently unchanged</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>+</td>
<td>Nonpermanently unchanged</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>?</td>
<td>Nonpermanently unchanged</td>
</tr>
</tbody>
</table>

The adjectives permanent and nonpermanent are used loosely to give a general indication of the groupings and should not be interpreted literally. For brevity, we may write temporary in place of nonpermanent. Categories 1 and 2 are comprised of spending units who expect to continue changing in the same direction or remain on a level position. Con-

\(^1\) This group consists of N.A.'s and of respondents who answered that they did not know what to expect.
sidering a year as a long time, this type of change is called permanent. Categories 3 and 4 are comprised of spending units who expect a quick reversal of recent income changes or who do not know what to expect. The latter subcategory is called temporary only in a potential sense. Manifestly it would be preferable to split the first four categories into eight separate categories, but we are forced by the small size of our sample to make the indicated combinations. Even the frequencies in the combined categories are not all as large as we should like for sound statistical treatment. Categories 5, 6, 7, and 8 are listed for completeness but consist of so few observations that no reliable statistical conclusions can be drawn from them.

There are various ways of trying to infer the relation between perception of income change and savings; we have begun a study of attitudinal variables by examining them in relation to the estimates of \( u_1 \) and \( u_2 \) in equations (1) and (2). The sense of this procedure is that we first extract the a priori suspected and measurable effects of income, family size, liquid assets, income change, and age on savings. Income change, the variable under comprehensive investigation, has a numerically small coefficient in (1) and (2), so its presence or absence is not important. Age is more at the borderline of importance but should be used as an explanatory variable in the present context since the relative permanence of income change may be thought of as related to age. We try, insofar as it is possible within the framework of (1) and (2), to account for the separate influence of age and then examine the influence of the perceived permanence of income change.

Properties of the estimated values of \( u_1 \) and \( u_2 \) in (1) and (2) are:

\[
\text{Mean } u_1 = 0 = \bar{u}_1, \\
\text{Mean } u_2 = 0 = \bar{u}_2, \\
\text{Standard deviation of } u_1 = S_{u_1} = 0.42, \\
\text{Standard deviation of } u_2 = S_{u_2} = 0.56.
\]

The frequency distributions of \( u_1 \) and \( u_2 \) are both bell shaped and roughly symmetrical about 0 but show a significant departure from normality with excessive modal concentration (leptokurtosis). If the distributions were more closely related to the normal distribution, we should be tempted to argue that \( u_1 \) and \( u_2 \) show chance characteristics that can be explained by purely random processes. The exact nature of the frequency distributions observed does not exclude the possibility that they are random drawings from some nonnormal population, but it does cause

\[\text{It is worth pointing out that liquid assets are significantly influenced by age; therefore, some additional age influence works through } SL_{-1}/Y \text{ on } S/Y.\]
us to search for additional systematic influences that help to explain variations in $S/Y$.

An analysis of the residuals, $u_1$ and $u_2$, within income change and expected income change categories shows the results presented in the table below. These findings show marked differences between the effects of income increases and income decreases on savings and, within the group of income decreases, quite different behavior for those experiencing permanent decreases from that for those experiencing temporary decreases. The most straightforward interpretation is that spending units experiencing permanent decreases of income save more than they otherwise would, while those experiencing temporary decreases maintain their customary level of living and save less than they otherwise would, in anticipation of a stronger income position in the near future. Very little can be said about income increases of either type. The residuals, $u_i$, average out to values near zero for both groups of income increase but are dispersed around zero in such a fashion that we cannot say whether spending units save more or less under the impact of either type of income increase.

<table>
<thead>
<tr>
<th>Category</th>
<th>Income Change 1947-1948</th>
<th>Expected Income Change 1948-1949</th>
<th>Frequency</th>
<th>$\bar{u}_1$</th>
<th>Frequency</th>
<th>$\bar{u}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
<td>+ or 0</td>
<td>144</td>
<td>-0.026</td>
<td>152</td>
<td>-0.005</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>- or 0</td>
<td>52</td>
<td>0.157</td>
<td>49</td>
<td>0.196</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>- or ?</td>
<td>47</td>
<td>-0.028</td>
<td>50</td>
<td>-0.028</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>+ or ?</td>
<td>27</td>
<td>-0.132</td>
<td>51</td>
<td>-0.152</td>
</tr>
</tbody>
</table>

The reader may get an idea of the significance of the means in the preceding table by forming the statistic

$$t = \frac{|\bar{u}_{ij}|}{S_{\bar{u}_{ij}}}$$  \quad (i = 1, 2; j = 1, 2, 3, 4)

and applying standard significance tests. The subscript $i$ refers to the home-owner or renter group, and $j$ refers to the income change category. $S_{\bar{u}_{ij}}$, can be approximated from

$$S_{\bar{u}_{ij}}^2 = \frac{S_{u_i}^2}{N_{ij}}$$

in which $S_{u_i}$ is the estimated variance of $u_i$ given previously, and $N_{ij}$ is the frequency taken from the appropriate cell of the preceding table. It is assumed that the $u_i$ in any income change category are drawings from a population with variance $\sigma_{u_i}^2$, estimated by $S_{u_i}$, and unknown mean, estimated by $\bar{u}_{ij}$. The hypothesis that the $u_i$ are
normally distributed was rejected; this does not, however, imply that the \( \bar{u}_x \) are nonnormal variates. The \( t \)-test is thus indicative of significance but not conclusive.

Since the means, \( \bar{u}_x \), are weighted means, the formula gives the minimum value of \( S_{\bar{u}_x} \), that value which would result if all weights were equal. However, equation (9) does provide a very close approximation, which is entirely adequate for our purposes. Calculations of the appropriately weighted variances in many cases show that the minimum estimate in (9) is only negligibly lower than the correct value.

Calculation of the statistic

\[
\begin{equation}
\begin{gathered}
t = \left| \frac{\bar{u}_{ij} - \bar{u}_{ij+1}}{S_{\bar{u}_x, ij+1}} \right| \\
(i = 1, 2; j = 1, 3),
\end{gathered}
\end{equation}
\]

in which

\[
\begin{equation}
S_{\bar{u}_x, ij+1} = \sqrt{S_{\bar{u}_x}^2/N_0} + \sqrt{S_{\bar{u}_x}^2/N_{ij+1}},
\end{equation}
\]

indicates that \( \bar{u}_1 \) is significantly less than \( \bar{u}_2 \) and that \( \bar{u}_3 \) is near the borderline of being significantly greater than \( \bar{u}_4 \). Both tails of the \( t \)-distribution are used in judging the significance of these differences, but if we make one-sided tests against only positive or negative alternatives, as the case may be, all differences are significant. The economic hypotheses implied here are that permanent income change is negatively related to savings and that temporary income change is positively related to savings. Further examination of these hypotheses can be made by dividing the sample into two groups, those with permanent and those with temporary income change. Equations of the form (3) are then estimated for each group separately. If the coefficient of \( (Y - Y_1)/Y_1 \) is definitely negative for the permanent group and definitely positive for the temporary group, then we have further evidence in support of our hypotheses. Calculations do not show a reliable coefficient of income change in the newly estimated equation for the permanent group; therefore we cannot, at this stage, say more than that income decrease affects saving in two different ways, depending on whether the decrease is permanent or temporary, and that income increase of either type may affect savings positively or negatively, although the average effect in our sample is near zero.

Income expectations were used in the preceding analysis to indicate the way respondents perceived their experienced income change, but expectations usually are introduced in theoretical economics in advance of decision-making and are used as a predetermined variable to explain actions taken at a later time point. From the 1948 Survey of Consumer Finances we can obtain income expectations, as of January–February 1948, from spending units whose savings behavior during the calendar year is later reported in the 1949 Survey of Consumer Finances. Let us
first consider the four groups: (1) those units who expected their income to rise from 1947 to 1948, (2) those who expected their income to remain about the same, (3) those who expected their income to fall, and (4) those whose income expectation is unknown. The mean residuals from the savings functions within each class are given in the accompanying table. The only remarkable figures in this calculation are the two substantial positive entries in the third row, those expecting an income decline, and negative entries in the fourth row, those whose expectation is not known. The entries in the third and fourth rows are not, however, significantly different from zero at the 5% level of significance when either one or two tails of the t-distribution are used. The results for those expecting an income decrease are more convincing, however, since they fit in better with a priori reasoning and can be more definitely classified with respect to the variable under consideration. We are presented with the tentative hypothesis that expected income drops induce preparatory savings, ceteris paribus. The findings for the other groups in the table are uninformative.

<table>
<thead>
<tr>
<th>Expected Income Change 1947–1948</th>
<th>Frequency</th>
<th>$\bar{u}_1$</th>
<th>Frequency</th>
<th>$\bar{u}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>60</td>
<td>-0.012</td>
<td>92</td>
<td>0.031</td>
</tr>
<tr>
<td>0</td>
<td>151</td>
<td>0.017</td>
<td>151</td>
<td>-0.025</td>
</tr>
<tr>
<td>-</td>
<td>47</td>
<td>0.068</td>
<td>40</td>
<td>0.110</td>
</tr>
<tr>
<td>?</td>
<td>30</td>
<td>-0.075</td>
<td>35</td>
<td>-0.054</td>
</tr>
</tbody>
</table>

There is a limit to the further cross classification of expected income change within our small sample, but we may go as far as we did in the previous example and make a joint classification with realized income change. Some of the subdivisions of the expected income groupings with respect to realized income increase, +, and income decrease, −, are given in the table below.

<table>
<thead>
<tr>
<th>Expected Income Change 1947–1948</th>
<th>Income Change 1947–1948</th>
<th>Frequency</th>
<th>$\bar{u}_1$</th>
<th>Frequency</th>
<th>$\bar{u}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>50</td>
<td>-0.024</td>
<td>68</td>
<td>0.030</td>
</tr>
<tr>
<td>0 or −</td>
<td>+</td>
<td>120</td>
<td>-0.015</td>
<td>112</td>
<td>-0.028</td>
</tr>
<tr>
<td>+ or 0</td>
<td>−</td>
<td>50</td>
<td>0.070</td>
<td>68</td>
<td>-0.002</td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>22</td>
<td>0.074</td>
<td>21</td>
<td>0.133</td>
</tr>
<tr>
<td>?</td>
<td>+</td>
<td>21</td>
<td>-0.089</td>
<td>22</td>
<td>-0.059</td>
</tr>
</tbody>
</table>

\[\text{As in the previous case, this group consists of N.A.'s and of respondents who answered that they did not know what to expect.}\]
These classifications try to pick out those who expected increases and received them, those who expected no change or decreases and got increases, those who expected increases or no change and got decreases, those who expected decreases and received them, and those whose expectations were not determined and who got increases. We find no systematic patterns in the first three groupings but possibly can sharpen the previous findings by showing that, among the households stimulated to save in the face of expected income declines, this behavior was most marked among those actually experiencing the declines and that, among the households stimulated not to save in the face of uncertain or unknown income expectations, this behavior was most marked among those actually experiencing increases. The latter two findings are based on extremely small frequencies and are not rigorously reliable.

The approach we have used may seem inefficient because the income expectation and income change classes are too broad. In the calculations underlying equations (1)–(7), dollar amounts of income change have been used, but it is extremely difficult to obtain data on income expectations in dollar amounts or even in brackets of $1,000 or $2,000. In the 1948 Survey of Consumer Finances, respondents were asked to give income expectations in one of eight brackets, but the results were disappointing in that more than a third of the sample had to be coded as N.A. On the other hand, the response was much better to a broader question asking for expected increase, constancy, or decrease. An analysis of the bracket data similar to that carried out above for the four simple categories of expectations produces no useful information, although we are able to draw some tentative conclusions from the rougher data. This should serve as a warning to econometricians not to try to overwork the ability of respondents to answer questions in a typical interview. Naturally, we can expect the potentialities of survey methods to grow in the future. We must also search for techniques of data analysis that can make the fullest use of the limited types of information available.

It is instructive to examine some attitudinal data that refer to more general magnitudes than a respondent's personal variables. At the beginning of 1948 a question was posed to respondents on their general economic outlook for the entire country during the coming year.\footnote{The question was worded, "Do you think we will have good times or bad times or what during the next twelve months or so?"} We apply, again loosely, the adjectives \textit{optimist} and \textit{pessimist} to the replies and see whether the different types of attitudes are systematically related to the \(u_t\).

Only five categories of attitude response are used in this calculation. They are: (1) spending units who expected good times to prevail during 1948, (2) units who expected neither bad nor good times to prevail, (3)
units who expected bad times to prevail, (4) units who did not know what to expect, and (5) units whose attitudes were not ascertained.

<table>
<thead>
<tr>
<th>General Economic Outlook</th>
<th>Frequency</th>
<th>$\hat{a}$</th>
<th>Frequency</th>
<th>$\hat{a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good times</td>
<td>150</td>
<td>-0.034</td>
<td>149</td>
<td>-0.036</td>
</tr>
<tr>
<td>Neither bad nor good</td>
<td>13</td>
<td>0.142</td>
<td>21</td>
<td>-0.032</td>
</tr>
<tr>
<td>Bad times</td>
<td>76</td>
<td>0.082</td>
<td>87</td>
<td>0.044</td>
</tr>
<tr>
<td>?</td>
<td>22</td>
<td>-0.069</td>
<td>40</td>
<td>0.081</td>
</tr>
<tr>
<td>N.A.</td>
<td>27</td>
<td>-0.013</td>
<td>21</td>
<td>0.004</td>
</tr>
</tbody>
</table>

On inspection of the results given in the accompanying table, we find one important pattern. For both renters and home-owners $\hat{a}$, is lower in the first row (the optimists) than in the third row (the pessimists). Nothing definite appears for the other categories, but the frequencies are small. The approximate formulas for testing the significance of differences indicates that the difference is just significant for the homeowner group but not for the renter group. However, the agreement in the results for the two groups plus the agreement with a priori conceptions of the influence of optimism or pessimism lead us to accept the findings tentatively as a measure of the particular attitudinal effect on savings behavior.

The analysis of the residual variation in equations (1) and (2) can be carried further to help pick out interaction in the structural relations. The fact that liquid assets have a different effect on savings for spending units with income decreases from the effect on savings for those with income increases has already been discussed. Income change, however, is related to the income level, and it is possible that the appropriate interaction is between liquid assets and income rather than between liquid assets and income change. In an equation of the form

$$ S/Y = \beta_0 + \beta_1 \log Y/N + \beta_2 L_{-1}/Y $$

$$ + \beta_3 (Y - Y_{-1})/Y_{-1} + \beta_4 a + \beta_5 L_{-1} + u, $$

the marginal effect of liquid assets on savings is

$$ \frac{\partial S}{\partial L_{-1}} = \beta_3 + \beta_4 Y. $$

Equation (13) states that the marginal asset effect depends on the income level, which is analogous to the previous results which claimed that the marginal asset effect depends on income change. Either case may be considered as a representation of an interaction effect.
Mean values of \( u_1 \) in four different liquid asset classes are shown in the accompanying table. Except for the low value of \( u_1 \) in the $2,000-$4,999 bracket, these calculations indicate a positive value for \( \beta_0 \) in an equation of the form (12).

<table>
<thead>
<tr>
<th>Liquid Assets ( L_{-1} )</th>
<th>Frequency</th>
<th>( u_1 )</th>
<th>Frequency</th>
<th>( u_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ 0-$499</td>
<td>108</td>
<td>-0.042</td>
<td>156</td>
<td>-0.021</td>
</tr>
<tr>
<td>$ 500-$1,999</td>
<td>79</td>
<td>-0.059</td>
<td>91</td>
<td>0.006</td>
</tr>
<tr>
<td>$2,000-$4,999</td>
<td>55</td>
<td>0.191</td>
<td>44</td>
<td>-0.144</td>
</tr>
<tr>
<td>$5,000 and over</td>
<td>46</td>
<td></td>
<td>27</td>
<td>0.308</td>
</tr>
</tbody>
</table>

Equation (12) has been estimated from the present reinterview sample with the addition of \( H \) and log \( N \) as separate variables. The previous empirical equations are written essentially in per capita form. We now consider the possibility that spending unit size has an independent effect from that implied by relations among per capita variables.

Home-owners and renters (606 cases)

\[
S/Y = -0.65 + 0.21 \log Y/N - 0.12 \log N - 0.29 L_{-1}/Y \\
+ 0.06 (Y - Y_{-1})/Y_{-1} + 0.0026a + 0.000039 L_{-1} + 0.11H + u_4, \\
(0.09) \quad (0.11) \quad (0.02) \quad (0.04) \quad (0.0017) \quad (0.000006) \quad (0.05)
\]

\[
R = 0.55, \quad S_a = 0.49.
\]

Equation (14) confirms the indications of the preceding table that the marginal asset effect depends on the income level. A major difference between equations (14) and (3) is that the curvature of the savings-income relation has been reduced considerably by the use of \( L_{-1} \) as a variable in the former equation. This can be seen by the drop in the coefficient of \( \log Y \), the term mainly responsible for the curvature. The coefficient of \( \log N \) in (14) is not significant, indicating that the departure from a per capita formulation is not important. The same result follows whether or not \( L_{-1} \), the only spending unit aggregate, appears in the equation. If we combine the two coefficients of \( \log N \) in (14) to get -0.33, we find the standard error of the combined coefficient to be approximately 0.1, showing that spending unit size does play a significant role in the savings equation.

A serious limitation of attempts to infer econometric relations from observations gathered in a single survey is that prices and other market variables are held nearly constant during a short interview period. Time series aggregates or time series of cross-section samples are needed to
estimate many of the parameters associated with prices. Although prices paid do not vary during a snapshot picture of the economy,\textsuperscript{13} price expectations do vary from household to household. The influence of price expectations or even prices paid on household behavior is a dubious matter, but one must study the available data more intensively to draw definitive conclusions about this variable. In the survey conducted during January–February, 1948, price expectations were asked of respondents, and it would seem reasonable to study the responses in relation to the $u_t$ in the same way that we studied other attitudes. Unfortunately, there was a sharp break in basic commodity prices in the middle of the interviewing period so that responses refer to two different price situations. Another difficulty in studying price anticipation within the present context is that the savings concept is too aggregative. We are not sure in advance whether our savings data should vary inversely or directly with price anticipations. Expenditures on real estate and purchases of equity shares are commonly viewed as inflation hedges and would be expected to vary in the same direction as price anticipations. Other types of savings, such as changes in bank accounts, purchases of bonds, or purchases of insurance, are desirable channels for funds under falling prices but not under rising prices. We have presented these possible cross currents merely to show the problem involved. We have the complication involved in the collection of the particular set of price expectations and the problem of formulating the a priori hypothesis. The data are not informative either; they show no systematic or significant relation to the $u_t$. The present writer knows of no other set of survey data which allows one to study beginning-of-period price expectations and savings within a multivariate relationship, but this is a problem well worth investigating in future surveys.

The preceding examples show the way in which econometric analysis can proceed with survey data and especially how new types of variables can be introduced in the process of refining the random disturbances to a point at which we can have more confidence in their purely stochastic character. Other variables of a demographic nature, such as changes in family composition, race,\textsuperscript{14} and education have been studied in the

\textsuperscript{13} Geographical price differentials do exist, and Mendershausen, in an earlier study (Horet Mendershausen, "Differences in Family Savings between Cities of Different Size and Location, Whites and Negroes," \textit{Review of Economic Statistics}, Vol. 22, August, 1940, pp. 122–137), did attempt to measure the influence of price level on savings behavior. He did not find conclusive evidence of price effects, but the data on intertemporal price variations are rather unsatisfactory.

\textsuperscript{14} There are too few nonwhites in the reinterview sample to draw any reliable conclusions about racial influences, but our tabulations do show, in both homeowner and renter groups, positive values of $u_t$ for nonwhites. Mendershausen, \textit{ibid.}, reports higher savings at given income levels for Negroes than for Whites.
same manner, although no striking empirical results have yet been obtained with these variables. Eventually this process should lead to much better estimates of the underlying structural behavior patterns than we now have.

The annual Surveys of Consumer Finances place us in a much richer position for analyzing the structure of household behavior, but there are some definite improvements that are well within the realm of possibility. The small sample size used in these surveys prevents us from studying a certain type of phenomenon that is suggested by many sources, namely, the phenomenon of social group behavior. We may argue that households identify themselves as members of socioeconomic groups and make their economic decisions, to a certain extent, in relation to group means or norms.

A study by Brady and Friedman\textsuperscript{16} presents us with a specific hypothesis of this type and some empirical evidence. Their savings relation makes the savings ratio a function of the order position of income in a distribution within a group. More specifically, they make the savings ratio of an income percentile a function of the logarithm of the percentile. The income percentiles are chosen from distributions within a fixed geographical region. In order to study this aspect of savings behavior, the sample must provide reliable income distributions or mean incomes of the relevant social groups. One cannot go far in this respect with samples of 3,500. It is necessary to have much larger samples or pinpoint samples of moderate size restricted to well-defined socioeconomic groups.

The existing large samples from which Brady and Friedman made their calculations are inadequate because they do not provide asset data, attitudinal data, or information from which to select the relevant social groups. Purely geographical groupings will not serve the purpose.

Surveys of the type exemplified by the sample we have used in this paper will become both cumbersome and expensive if conducted for a large sample of 25,000 or more households. The alternative would be to proceed more modestly with a series of smaller surveys directed at getting all the necessary variables. If the problem of group influence were to be studied, this would mean careful advance limitation of the universe sampled, and little attempt should be made, in the first instance, to generalize about household behavior of the nation as a whole.

Two economic matters which need to be advanced in future surveys are these: (1) A set of concepts appropriate to individual households must be developed. (2) A complete accounting statement that has

elegance, precision, and full coverage needs to be developed where it is possible to obtain survey data. The economic concepts of household variables used in empirical work are largely analogues of entries in the macroeconomic social accounts and, in many cases, are ill-suited for microeconomic analysis. Some particular items that may well be treated differently at the two levels are capital gains, transfers, gifts, and inheritances.

When we are in a position to draw up a comprehensive set of household accounts on the basis of survey data, econometric analysis of the underlying relationships will best proceed with new developments. At present we are faced with too many gaps in trying to close the accounting system.

*National Bureau of Economic Research and Survey Research Center*