

DAVID BUNTING

The consumption function “paradox”

Introduction

One of the great insights of the *General Theory* was the idea of the consumption function. According to Keynes (1936, p. 96), consumption was related to income by a “fundamental psychological law”

upon which we are entitled to depend with great confidence both a priori from our knowledge of human nature and from detailed facts of experience, . . . that men are disposed, as a rule and on the average, to increase their consumption as their income increases, but not by as much as the increase in their income.

While simple in formulation, in application the law was soon enveloped in ambiguity. Empirical studies of consumption seemed to suggest more than one pattern of spending behavior. Friedman (1957), in his classic *A Theory of the Consumption Function*, cited estimates by Kuznets, Ferber, Goldsmith, and others of a time series or long-run marginal propensity to consume (MPC) around .90. On the other hand, studies utilizing expenditure survey data found cross-sectional or household MPCs mostly in the range of .60 to .80. How was it possible

The author is Professor of Economics at Eastern Washington University in Cheney, Washington.

Initial research for this project was funded by a grant from the Northwest Institute for Advanced Study. I am grateful to Peter Lin of the Washington State Department of Revenue and to members of the Economics Department Colloquium, especially Hadi Dadashi and Tom Karier, for their many comments and alternatives as the paper went through various versions.

A statistical appendix describing the various time series and cross-sectional datasets used is available on request from the author.

for spending behavior in the long run to be so much different than that in the cross section?

Another inconsistency arose in regard to the income elasticity of consumption ($Ncy = [dC/dY][Y/C]$). Since 1900 the average propensity to consume (APC) has remained near .90, implying a long-run income elasticity of about one and a constant savings rate. However, for cross-sectional behavior, the elasticity was much less than one, indicating a rising saving rate. Friedman (1957, p. 44) found this inconsistency fatal to Keynes' notion of a consumption function, at least for household behavior: "It is this feature, of course, that makes it impossible to regard these regressions as estimates of a stable relation between consumption and income."¹

Because of these empirical differences, the idea of a "paradox" arose—that somehow individual or household behavior was different than aggregate behavior. This paradox had profound implications. It implied that Keynes' consumption theory, and by extension the rest of the *General Theory*, was incomplete, possibly even incorrect. Consumption was fundamental to Keynesian economics, yet Keynes' theory produced inconsistencies and ambiguities. Failure on this point naturally cast suspicion on others. The paradox greatly dampened the policy implications of the *General Theory*. Cross-sectional studies dealt with household spending behavior at discrete points in time. These suggested much lower investment multipliers than deduced from long-run aggregate studies. Lower multipliers meant that fiscal policies to alter national income would be either less effective or more costly.

The paradox also stimulated efforts to devise a more complete theory of consumption. Of these, probably the best known are Duesenberry's (1949) relative income hypothesis, Friedman's (1957) permanent income hypothesis, and Modigliani and Brumberg's (1986, reviewing the literature) life-cycle hypothesis. As these efforts stimulated further ones and as the paradox persisted, Keynes's original formulation came to be considered more as an important first step than an accurate description of aggregate consumption.

The paradox reconsidered

It is important to remember that modern alternatives to Keynes's theory of consumption were originally devised to explain an empirical

¹In a later study of savings behavior alone, David and Scadding (1974, p. 238) found a relatively constant savings rate between 1900 and 1970.

result. While cast in the mode of deductive theorizing, both Friedman (1957, p. 4) and Modigliani (1986, p. 298) clearly indicated their theories were initially motivated by factual rather than theoretical considerations. Studies of household spending had found changes in current consumption were rather weakly influenced by changes in current income. Since current income was an unsatisfactory explanatory factor, perhaps some longer term measure of household income such as past peak income or average income over a period of years or even income over a lifetime might show stronger influence. But regardless of specifics, all the proposed theories depended upon the "fact" that changes in current consumption were much less than changes in current income, that is, the marginal propensity to consume was smaller than expected.

There is an obvious danger in building theories to explain a certain fact—the fact could change or disappear. Epicycles were devised to explain the rotation of stars around the Earth; bleeding was used to relieve headaches; some children ascribe large increases in wealth to the tooth fairy; Columbus' proposal seemed ridiculous to those who believed the world flat. Similarly, modern alternatives to Keynes' consumption function might seem plausible when household *MPCs* are .70 but could be irrelevant when the values are .90.

The "fact" of the consumption function paradox has been little questioned since Friedman's (1957) impressive review of thirty years ago. Since then both computational technologies and data sources have dramatically improved while new generations of economists have had the opportunity to mull over past learning. Also, the basic data used in consumption function research is uncomplicated. When the functions are re-estimated using old data, spending patterns still remain significantly different in the cross section as compared to the long run. However, these differences now seem more the result of statistical considerations than fundamental differences in economic behavior.

For example, all cross-sectional studies involve household or family spending behavior while the long-run studies deal with aggregate behavior. For valid comparisons of either marginal propensities or income elasticities, the spending units should be the same. Comparison should be on the basis of household *or* aggregate spending, not household *with* aggregate spending. While much has been made of the inconsistency of the cross-sectional studies' results with other results of consumption behavior, these studies are suspect on a number of points. As shown below, they sometimes produce incredible values for marginal or average consumption. Finally, the cross-sectional data is highly aggregated with thousands of households collapsed into a dozen or so

continuous but arbitrarily defined income categories. If biased, this grouping could fatally contaminate any study of cross-sectional behavior. Important spending characteristics could be hidden by aggregation, or regression coefficients derived from relatively few observations could be sensitive to grouping criteria.

Time series functions

Because they have been consistent with external evidence and theoretical expectations, long-run estimates of spending behavior such as Equation 1 are used as the benchmark in consumption function research. This function was estimated with deflated annual data taken from the 1929–1982 *National Income and Product Accounts* (1986). Variable definitions follow those outlined by Friedman (1957, pp. 42–3) and his primary data source Brady (1956, p. 181): income as personal income includes all income received while consumption includes personal expenditures, gifts, personal taxes, and nontax payments.² With *t*-statistics in parentheses, the long-run marginal propensity of .93 and income elasticity of 1.00 are commonplace results.

$$(1) \quad C = 1.93 + .9311Y$$

$$(0.14) \quad (105.34) \quad Ncy = 1.00$$

However, since the function describes aggregate behavior, it cannot be compared with cross-sectional results which involve family or household behavior. A valid comparison can be made by converting the time series data to a household basis by dividing both annual consumption (C_t) and income (Y_t) by the number of annual households (H_t). The re-estimated function then becomes:

$$(2) \quad CH = 1549.6 + .8699YH$$

$$(2.39) \quad (34.38) \quad Ncy = .93$$

where $CH_i = C_i/H_i$ and
 $YH_i = Y_i/H_i$.

The long-run *MPC* falls from .93 to .87 merely by placing the data in its correct dimension. This puts the consumption function paradox in a

²A *Statistical Appendix* detailing both data sources and variable definitions used in this paper is available upon request.

somewhat different perspective. Previously cross-sectional *MPCs* in the .60 to .80 range were compared against a standard of .90; now the standard is somewhat less. This suggests that differences in the two types of *MPCs* might be explained by circumstances such as data quality or statistical procedures rather than by behavioral factors.

The reason why the aggregate household data produced a different marginal propensity to consume estimate than the untransformed data exposes an important implicit assumption which underlies much consumption function thinking. That is, one dataset is assumed to be a linear transformation of the other. If this is correct, then conversion of the data from one form to another will have no effect on the value of the estimated *MPC*.

One of the elementary rules of the "algebra of expected values and sample statistics" is that the "variance of a variable multiplied by a constant is equal to the square of the constant times the variance of the variable" (Koutsoyiannis, 1977, pp. 539-540). That is,

$$(3) \quad \text{Var}(a+bY) = b^2\text{Var}(Y)$$

This is also true for covariances,

$$(4) \quad \text{Cov}(a+bC,a+bY) = b^2\text{Cov}(C,Y)$$

Since the slope of the regression of consumption on income is given by

$$(5) \quad \text{MPC} = \frac{\text{Cov}(C,Y)}{\text{Var}(Y)},$$

conversion of the aggregate data to a household basis will yield the same *MPCs* only if the transformation is linear; that is, if

$$(6) \quad CH = a + bC \text{ and } YH = a + bY.$$

If these conditions exist, then

$$(7) \quad \text{MPC} = \frac{\text{Cov}(C,Y)}{\text{VAR}(Y)} = \frac{b^2\text{Cov}(CH,YH)}{b^2\text{Var}(YH)}$$

In the case at hand, the differing *MPCs* indicate the transformation of C_t to CH_t and Y_t to YH_t is not linear. This is also indicated by the correlation of the variables in their two forms: $r(C, CH) = .9593$ and $r(Y, YH) = .9341$. Reasons for the nonlinear transformation are not hard to find. Since 1929 the growth in households has not exactly matched the growth in income. Households have declined in size, primarily reflecting a declining birth rate, and increased in relative numbers, reflecting changing marriage and living circumstances.

Cross-sectional functions

Estimates of cross-sectional consumption functions such as those found in Table 1 are usually derived from BLS consumer expenditure survey data. Nine of these surveys have been undertaken between 1888 and 1972.³ Because some included alternative definitions of spending units or income, thirteen household functions were estimated with eight of these producing *MPCs* in the .60 to .80 range. The income elasticities of consumption (N_{cy}) are all much less than one, suggesting a historic rising saving rate. Also the coefficients show no pattern which indicates changing spending behavior over time. In 1901 households had an *MPC* of .74 while in 1960 they had one of .75. Overall, the results are consistent with other reported examinations of cross-sectional household behavior.

However, closer inspection reveals some peculiarities. Despite alleged constancy, the *APCs* show surprising variation between 1888 and 1972, ranging from a high of .97 to a low of .47. This could suggest something is amiss with the data. The *MPCs* show considerable variability with some unreasonably low values. The 1930s were supposed to be terrible times for farmers. How could farm families in 1935, even those not on relief, have an *MPC* of .22 or, contrarily, an *MPS* of .78? The coefficients seem sensitive to alternative household or income definitions. *MPCs* from the 1935 survey varied by twofold from one household type to another while different income definitions in 1944 and 1972 produced different *MPCs*.⁴ Finally, although each BLS survey usually involved thousands of respondents selected to represent some larger segment of all households, the data were never organized

³Results for the 1980-81 survey were not considered because they were reported in much different form than for previous ones.

⁴In the 1944 survey, sample b included war bonds as consumption; in the 1972 survey, sample b included more income groups.

Table 1

Estimates of Cross Sectional Consumption Functions ($C = a + bY$)

Survey Year	Intercept*	MPC*	APC	Ncy	N	Sample Description
1888	267.9	.51	.73	.70	12	Families
1901	125.5	.74	.91	.81	12	Normal families
1917	181.1	.79	.90	.88	7	Families
1935a	1,095.4	.22	.47	.48	9	Farm Nonrelief families
1935b	923.1	.49	.63	.78	15	All families
1935c	845.5	.54	.71	.76	9	Nonrelief, nonfarm families
1941	545.5	.71	.85	.84	9	Urban families
1944a	720.4	.66	.89	.74	9	Families
1944b	607.8	.78	.97	.80	9	Families, version 2
1950	1,109.2	.75	.93	.81	9	Urban families
1960	1,404.0	.75	.93	.81	10	Urban, rural families
1972a	2,696.8	.65	.88	.74	12	Households
1972b	3,869.3	.54	.73	.74	16	Households, version 2

*Significant at 99%.

Table 2

Average Consumption and Income: 1960 BLS Consumer Expenditure Survey

Income Level	Average Consumption	Average Income	APC	Percent Households
under 1000	1,212	691	1.75	3.5%
1000 to 2000	1,849	1,514	1.22	9.7
2000 to 3000	2,841	2,504	1.13	10.0
3000 to 4000	3,778	3,500	1.08	10.2
4000 to 5000	4,700	4,523	1.04	10.8
5000 to 6000	5,593	5,498	1.02	11.8
6000 to 7500	6,524	6,706	.97	14.4
7500 to 10000	8,067	8,569	.94	15.7
10000 to 15000	10,672	11,738	.91	10.4
15000 and over	18,299	22,584	.81	3.7

into more than sixteen income categories. This organization, if not a linear transformation of the original ungrouped data, could explain the peculiarities just indicated.

Nearly all the concerns regarding estimates of household consumption behavior can be illustrated using a complete, ungrouped BLS survey dataset. Of the two modern annual datasets available, a variant of the 1960 survey was selected.⁵ This sample, consisting of 13,694 rural and urban consumer units representing the spending behavior of more than 55 million households, was aggregated following BLS procedures.⁶ First, the data were grouped by arbitrary but seemingly reasonable income intervals such as "under \$1000," "\$1000 to \$2000," "\$2000 to \$3000," and so on. To show "typical" behavior, average income (YA) and average consumption (CA) were calculated by dividing each group by its number of households.

The results of this grouping and averaging are found in Table 2. Each table row contains the average of consumption and income included in the indicated income interval. Data in this form is that which has been used to estimate nearly all household consumption functions. In the case at hand, the estimate produces a classic "paradoxical" cross-sectional consumption: the MPC is lower than that for the long run while the Ncy suggests an increasing savings rate as income rises.

$$(8) \quad CA = 1035.9 + .7840YA$$

(6.53) (45.25) $Ncy = .84$

Does the estimated equation accurately describe cross-sectional behavior? This question can be considered with information always provided by the BLS with its published grouped data or easily extracted from the data itself. This information, also found in Table 2, is the average propensity to consume by income level, and the relative distribution of households, calculated from group totals.

Since ordinary least squares regressions implicitly weight each pair of observations equally, the grouping of the $APCs$ around 1.0 with extreme values at each end of the income distribution hints that the regression of average consumption on average income might be strongly influenced by extreme cases of savings and dissavings. Further, the percentage distribution of households clearly shows that some intervals represent far more households than others. For example, 3.5 percent of all households are represented by the "under \$1,000" interval while

⁵A description of the BLS dataset is found as an Appendix.

⁶A total of 34 households with negative incomes were omitted from the dataset to avoid understating average income in the under 1,000 group.

15.7 percent (more than four times the first group) are in the "\$7,500 to \$10,000" interval. More importantly, the relative distribution of households by income interval clearly indicates that the grouped data is not a linear transformation of the ungrouped data.

All but the last point have been made by Mayer (1972, p. 63). He objected to the grouping of data: "surely an income class with say, 20 percent of the population should receive a bigger weight than another containing only, say 5 percent." He also argued "there is no reason why the basic unit observation should be an income class rather than a household." Mayer thought the problem of unequal group size could be corrected by using a weighted regression to estimate behavior with the weights being group frequency. Although, as Equation 9 shows, this procedure slightly raises both the *MPC* and *Ncy*, it ignores the basic problem that the grouped data is not a linear transformation of the ungrouped data. A second resolution of the unequal group size problem has been to exclude the lowest and highest income groups. While reducing extremes in the dataset by removing both the poorest and richest households, there is no apparent theoretical justification for this procedure.⁷

$$(9) \quad CA = 1005.2 + .8028YA \quad Ncy = .86$$

(6.46) (39.62)

When the equation is re-estimated using the ungrouped dataset of 13,694 observations the cross-sectional function becomes:

$$(10) \quad CH = 796.04 + .8362YH \quad Ncy = .87$$

(32.07) (295.62)

Empirical differences between cross-sectional and long-run household spending behavior now virtually disappear. The *MPCs* from both types of functions are .84 and .87 while the *Ncys* are .87 and .93. Rather than developing some elaborate theory of consumption behavior to explain these differences, they can be explained by the quality of the data used.

⁷Extreme income groups are usually removed because they are disproportionately small compared to other groups in the sample. However, in some surveys, the extremes are disproportionately large. In the 1972 survey, the smallest income group represented 13.4 percent of the sample. For 1935 farm nonrelief families, the smallest group represented 17.7 percent of the survey. In the 1944 survey, the largest group represented 14.4 percent of the sample. Ad hoc exclusion of these groups seems to be a strong argument to use ungrouped rather than grouped data.

The long-run consumption function most likely overstates spending behavior because the "current expenditures with capital consumption allowance" of insurance companies, professional organizations, museums, philanthropic and eleemosynary associations, and similar organizations are included in the national accounts as personal expenditures.⁸ On the other hand, the cross-sectional functions probably understate consumption because the BLS survey respondents do not perfectly recall their expenditures.⁹ The net effect of these probable errors in the data is to reduce further the empirical differences between cross-sectional and long-run spending behavior.

While these findings strongly suggest that the consumption function paradox is simply the result of incorrect statistical procedures, the argument would be more convincing if additional cross-sectional datasets were examined. Research is currently underway using data collected for the 1980-81 BLS quarterly consumer expenditure survey. To date, a preliminary dataset involving 1,407 consumer units for the first quarter of 1980 has been constructed. This produced the following estimate:

$$(11) \quad CH = 1289.20 + .8225YH \quad Ncy = .70$$

(10.16) (40.27)

These results must be considered only provisional. Although the *MPC* of .82 is similar to that from the 1960 data, the *Ncy* of .70 is much too low. The cause of this is an *APC* of 1.17, an historically unprecedented magnitude. Until this quarterly estimate is validated by comparisons with others from other quarters as well as with a yearly estimate, Equation 11 must be considered only suggestive.

Conclusions

It seems incredible that two generations of economists could endlessly diagram and discuss the consumption function "paradox" without

⁸See U.S. Bureau of Economic Analysis (1986, p. 112) footnotes 13 through 18, 21, 22, 25, and 26.

⁹While the BLS does not indicate completeness for the 1960 survey, it estimates that its 1980-81 quarterly interview survey collected detailed data on an "estimated 60 to 70 percent of total household expenditures" and that other estimates "account for an additional 20 to 25 percent of total expenditures." See "1980-1981 Interview Survey Public Use Tape Documentation" which accompanies the 1980-81 Survey data tape.

ever considering the basic factors underlying the paradox itself. Yet this is exactly what happened. Thousands of theorists innocently drew diagrams where one line representing aggregate long-run spending crossed another representing household cross-sectional spending and never considered the units of measurement for their variables. Similarly, data conveniently compiled by the Bureau of Labor Statistics was used to estimate consumption functions and no one thought to question the appropriateness of data collected to describe expenditure patterns, organized by income interval, and highly aggregated. Finally, even though the notion is an elementary proposition of statistics, the issue of the correct transformation of ungrouped to grouped data was never considered.

Some observations and implications can be drawn from these conclusions. Modern macroeconomic theories of consumption were initially motivated by the conflict between time series and cross-sectional empirical results. Does the eradication of the conflict imply that theories involving relative income, life-cycle savings, and permanent income are no longer valid? Or can these theories stand on their own merits despite the fact that changes in current consumption are largely explained by changes in current income? It should be realized that the latter situation, if true, does not disprove the modern alternatives to Keynesian consumption theory. As Friedman (1957, p. 231) argued in regard to his own theory:

The possibility of dispensing with the (existing theory) does not, of course, mean that (existing) empirical findings are in error, that the variables (found) related to consumer behavior are not related to it, any more than acceptance of the Copernican view rendered nonexistent the astronomical movements that it was necessary to introduce additional epicycles to explain. What it does mean is that these empirical relations can all be inferred from a much simpler structure, that they can all be regarded as manifestations in different guise of a single and simpler set of forces rather than as the result of largely irreducible ultimate forces.

On another point, Friedman also argued that his particular division of income into permanent and transitory components implied discretionary fiscal policy would be largely ineffective. His theory meant that "a much larger part of current consumption is interpreted as autonomous and a much smaller part as dependent on current income and hence, through the multiplier process, on investment. The result is a smaller investment multiplier, and an inherently more stable system"

(1957, p. 238). Now that Friedman's distinctions seem unnecessary, do larger than expected multipliers mandate resurrection of aggressive fiscal policies to stabilize an economic system less stable than previously assumed?

In conclusion, the results presented here indicate that Keynes's "fundamental law" is valid, that as a rule and on the average, households are disposed to increase their consumption as their income increases, but not by as much as the increase in their income. Obviously the merits of this conclusion depend upon further study and research. The BLS datasets are large and complex but relatively cheap; the BLS staff is helpful. Hopefully, these data and conflicting perspectives can produce a clearer understanding of aggregate consumption behavior.

Perhaps the most most useful way to view spending behavior is in terms of two separate consumption functions. One of these, based on grouped data organized by income interval, shows that on average people spend a smaller fraction of their income as their income increases. Representing the traditional tilted cross-sectional relationship of negative savings at low incomes and positive savings at high incomes, the major theoretical importance of this function is for predicting individual responses to income changes. The second function, based on either ungrouped cross-sectional data or long-run aggregate data, indicates the relationship between aggregate household consumption and income. The major theoretical importance of this function is for predicting aggregate responses to changes in aggregate income. Together, these functions offer a simpler but more complete explanation of spending behavior than that provided by one function alone.

APPENDIX

The 1960 Survey of Consumer Expenditures

A summary description of the survey with citations to more technical studies is found in U.S. National Technical Information Service (1980, p. 48). As described in the "Handbook for Survey Users," distributed to purchasers of the survey tape, the 1960-61 survey of consumer expenditures involved an elaborate national sample, stratified on the basis of geographic, urban city size, and rural housing characteristics. Ultimately 13,728 survey schedules were completed. Each of these was weighted by an "expansion factor," "the estimated consumer units in the universe represented by each completed schedule in a primary

sampling unit." The BLS cautioned that the sample was not self-weighting: "Aside from the various differential sampling rates for the urban population among individual SMSA's, unequal overall sampling rates were used for urban, rural nonfarm, and rural farm components."

For purposes of estimation, income, defined as "money income," included earnings, profits, interest, dividends, public unemployment and social security benefits, pensions, public assistance and private relief. Consumption included all "expenditures for current consumption" or "current living expenses," gifts and contributions, and personal taxes, including Federal, State, and local income and property taxes and excluded all "personal insurance" payments for life, endowment, and annuity insurance, including pensions, social security, and group plans. Finally, each item was multiplied by its appropriate expansion factor.

REFERENCES

- Brady, D.S. "Family Saving, 1898 to 1950." In Goldsmith, R.W., D.S. Brady, and H. Mendershausen, *A Study of Saving in the United States*, vol. III. Princeton, NJ: Princeton University Press, 1956.
- David, P.A., and Scadding, J.L. "Private Savings: Ultrarationality, Aggregation, and 'Denison's Law'." *Journal of Political Economy*, 1974, 82, 225-247.
- Duesenberry, J.S. *Income, Saving, and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press, 1949.
- Friedman, M. *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press, 1957.
- Keynes, J. M. *The General Theory of Employment Interest and Money*. New York: Harcourt, Brace & World, 1937.
- Koutsoyiannis, A. *Theory of Econometrics*, 2nd ed. New York: Barnes & Noble, 1977.
- Mayer, T. *Permanent Income, Wealth and Consumption*. Berkeley: University of California Press, 1972.
- Modigliani, F. "Life Cycle, Individual Thrift, and the Wealth of Nations." *American Economic Review*, 1986, 76, 297-313.
- U.S. Bureau of Economic Analysis. *The National Income and Product Accounts of the United States, 1929-82*. Washington DC: Government Printing Office, 1986.
- U.S. National Technical Information Service. "Survey of Consumer Expenditures—1960-61." *A Directory of Computerized Data Files and Related Technical Reports*. Washington DC: Government Printing Office, 1980.

