

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: A Theory of the Consumption Function

Volume Author/Editor: Milton Friedman

Volume Publisher: Princeton University Press

Volume ISBN: 0-691-04182-2

Volume URL: <http://www.nber.org/books/frie57-1>

Publication Date: 1957

Chapter Title: Evidence from Income Data on the Relative Importance of Permanent and Transitory Components of Income

Chapter Author: Milton Friedman

Chapter URL: <http://www.nber.org/chapters/c4409>

Chapter pages in book: (p. 183 - 199)

CHAPTER VII

Evidence from Income Data on the Relative Importance of Permanent and Transitory Components of Income

THE distinguishing feature of the permanent income hypothesis is the central role it assigns to certain characteristics of the income distribution in explaining empirical evidence on consumption behavior. In examining the consistency of our hypothesis with available evidence, we have so far restricted ourselves to inferring these characteristics from paired data on consumption and income; from data for either a number of consumer units in the same year, or a number of years for a group of consumer units. As has been noted several times, however, these characteristics can be determined from income data alone. An estimate of P_y , the fraction of the variance of incomes contributed by the permanent component, can be constructed from data on the incomes of identical consumer units in different years.¹ This possibility provides an independent means of testing our hypothesis; in addition, it enhances its potential usefulness by broadening the range of data that it can be used to interpret.

Consider a group of consumer units whose measured incomes we know for two successive years. Suppose that the differences among the incomes of the members of the group in each year are entirely attributable to differences in permanent components of income, and that we can neglect "aging" from one year to the next (alternatively, assume that all members of the group are affected alike by aging). The relative measured income position of the members of the group would then, in some sense yet to be defined precisely, be the same in the two years; there would be perfect correlation, also in a sense yet to be defined precisely, between their incomes in the two years. At the other extreme, suppose all differences in income among members of the group in at least one of the years are attributable to transitory

¹ Similarly, estimates of P_c , the fraction of the total variance of consumption expenditures contributed by the permanent component of consumption, can be constructed from data on the consumption expenditures of identical consumer units in different years. The discussion that follows about P_y applies with suitable changes equally to P_c .

EVIDENCE FROM INCOME DATA

factors that exhaust their effect in that year. Incomes in the second year would then tend to be uncorrelated with those in the first year. The size of the correlation between incomes in two successive years therefore provides some evidence on the importance of the permanent component in producing differences in measured income.

1. A Method of Estimating P_v

Let \bar{y}_i stand for the mean measured income and s_i , for the standard deviation of measured income in year i ; r_{ij} , for the product moment correlation coefficient between measured incomes in years i and j ; b_{ji} , for the slope of the computed regression of year j 's measured income on year i 's; b_{ij} for the slope of the regression of year i 's measured income on year j 's; and P_i for the fraction of the total variance of measured income contributed by the permanent component in year i . I have elsewhere suggested two alternative statistical estimates of P_i , differing in the precise meaning attached to constancy of the permanent component.² One is derived for what I have called the *mean assumption*, which is that permanent components maintain the same ratio to the mean of the group in different years, so that the relative variability of the permanent component is unchanged. Under this assumption,³

$$(7.1) \quad \begin{cases} P_i = b_{ji} \frac{\bar{x}_i}{\bar{x}_j} = r_{ij} \frac{s_j \bar{x}_i}{s_i \bar{x}_j}, \\ P_j = b_{ij} \frac{\bar{x}_j}{\bar{x}_i} = r_{ij} \frac{s_i \bar{x}_j}{s_j \bar{x}_i}. \end{cases}$$

The other measure is derived for what I have called the *variability assumption*, which is that the fraction of the total variability contributed by the permanent components is the same in successive years, i.e. that P_v is the same in years i and j . Let P' designate the common

² What follows is essentially a restatement and summary of Friedman and Kuznets, *Income from Independent Professional Practice*, pp. 325-338, 352-364, which contains full proofs of the formulas that follow.

³ Under the logarithmic variant, the mean assumption can be broadened without affecting the results. If the absolute values of the permanent components are in a common ratio, whether the ratio of the arithmetic means or any other common ratio, estimates of P_v are given by

$$(7.1) \quad \begin{cases} P_{Y_i} = B_{ji}, \\ P_{Y_j} = B_{ij}, \end{cases}$$

where B_{ji} is the regression coefficient of the regression of the logarithm of income in year j on the logarithm of income in year i , and B_{ij} the same with i and j interchanged.

EVIDENCE FROM INCOME DATA

value of P_y . Then, under this assumption,⁴

$$(7.2) \quad P' = \sqrt{P_i P_j} = r_{ij}.$$

It is obvious from the formula that the contribution of the permanent component as estimated under the variability assumption is a geometric average of the contributions in the two years as estimated under the mean assumption. Although the variability assumption leads to mathematically simpler results, it is less attractive theoretically than the mean assumption. Further, some statistical tests made by Margaret Reid suggest that the mean assumption yields the better results in interpreting consumption data.⁵ These formulas are for incomes in absolute units; however, very similar formulas apply when the data are expressed in logarithms, and the logarithmic variant in general seems to fit the empirical evidence rather better.

Consider three successive years, say years 1, 2, and 3. Suppose that two estimates of P_3 are computed by (7.1): first, from data for years 2 and 3 (call this $P_{3.2}$); second, from data for years 1 and 3 (call this $P_{3.1}$). Clearly, the two results need not be identical. In general $P_{3.2}$ can be expected to be larger than $P_{3.1}$ since the correlation between incomes in two consecutive years can be expected to be higher than in two nonconsecutive years with one year intervening. This difference in numerical results reflects an implicit difference in the definition of the permanent component—a point that we have mentioned at several points but have not hitherto had occasion to state precisely. In taking $P_{3.2}$ as an estimate of the fraction of variance contributed by the permanent component in year 3, we implicitly define the permanent component as the component that is attributable to factors affecting income alike in two or more successive years, and the corresponding transitory component, as the component that is attributable to factors affecting income in one and only one year. In taking $P_{3.1}$ as an estimate of P_3 , we implicitly define the permanent component as the component that is attributable to factors affecting income alike in three or more successive years, and the corresponding transitory component, as the component that is attributable to factors affecting income in one or two but not three successive years.

⁴ Under the logarithmic variant, the formula is essentially the same if the variability assumption is interpreted as meaning that the same fraction of the logarithmic variance is contributed by the permanent component in successive years, namely,

$$(7.2') \quad P'_Y = \sqrt{P_{Y_i} P_{Y_j}} = r_{Y_i Y_j}$$

⁵ The results are contained in an unpublished paper by Margaret Reid, entitled "The Relation of the Within-Group Transitory Component of Incomes to the Income Elasticity of Family Expenditures."

EVIDENCE FROM INCOME DATA

More generally, we can conceive of income as the sum of a continuum of components, classified by the length of the period for which the corresponding factors affect income and the time unit in which they make their first appearance. For simplicity, we can represent this continuum by a trichotomy of permanent, quasi-permanent, and transitory components, where the "truly" permanent component is attributed to factors affecting income over the longest period considered; the "truly" transitory component, to factors affecting income in only a single time unit; the quasi-permanent component, to other factors affecting income in more than one, but not all, time units. Given data on incomes in a series of years, statistical estimates can be constructed of the fraction of the total variance of income contributed by each component.⁶

As the permanent income hypothesis is used to interpret empirical data, it may be necessary to elaborate it by allowing for the separate influence of quasi-permanent components, or even additional sub-components.⁷ For the present, however, it seems better to stick to the simpler formulation in terms of permanent and transitory components alone; this still leaves considerable leeway in the precise definition of the permanent component, which, as noted earlier, should be determined empirically, not imposed a priori.⁸

The approach to the analysis of income data just described can be used to estimate not only the contribution of the various components to the variance of income for the group as a whole but also their contribution to the deviation of the average income of a particular income class from the average income of the group as a whole. If the fraction of this deviation contributed by the permanent component (when we classify all components of income into the dichotomy of permanent and transitory components) is the same for all income classes in year i , then the regression of year j income on year i income will be linear, and conversely. These are also the conditions, therefore, under which, on our hypothesis, the regression of consumption on income will be linear for year i .

It should be noted that this analysis applies only to differences in income; it does not, for any one group, give evidence on the mean transitory component of income. Some such evidence can be obtained, however, by applying the same analysis to data on the mean incomes of a number of groups, say to the mean incomes in two successive years of a set of communities.

⁶ See Friedman and Kuznets, *op. cit.*, pp. 352-64.

⁷ This is equivalent to allowing for different "short-run" and "long-run" marginal propensities to consume.

⁸ See above, pp. 23-25, 92-93, 142, 150-151.

EVIDENCE FROM INCOME DATA

TABLE 18
Correlation Coefficients between Incomes of Identical Units
in Different Years

<i>Group</i>	<i>Income Definition</i>	<i>Years Correlated</i>	<i>Number of Years Intervening between Years Correlated</i>	<i>Correlation Coefficient</i>
1. Physicians	Earnings from independent practice	a	0	.93
		b	1	.91
		1929 & 1932	2	.88
2. Dentists	Earnings from independent practice	a	0	.92
		b	1	.88
		1929 & 1932	2	.76
3. Lawyers	Earnings from independent practice	c 1932 & 1934	0 1	.84 .80
4. Certified public accountants	Earnings from independent practice	d	0	.88
		e	1	.78
		1929 & 1932	2	.82
5. Consulting engineers	Earnings from independent practice	1929 & 1930	0	.67
		1929 & 1931	1	.63
		1929 & 1932	2	.52
6. Families in 33 cities	Total family income	1929 & 1933	3	.74 ^t (.63 to .86)
7. Wisconsin taxpayers grouped in family unit	"Economic income" of family unit	g	0	.84
		h	1	.78
		i	2	.76
		j	3	.71
		k	4	.70
1929 & 1935	5	.69		
8. Urban spending units	Total income	1947 & 1948	0	.83
9. Farm families reporting to agricultural experiment station	Logarithm of family net cash income	l	0	.41
		1940 & 1942	1	.33 ^m
10. FHA families	Logarithm of family net cash income	l	0	.48
		1940 & 1942	1	.46 ^m
11. Groups 9 and 10	Logarithm of family net cash income	l	0	.52
		1940 & 1942	1	.47 ^m

(cont. on next page)

EVIDENCE FROM INCOME DATA

TABLE 18 (cont.)

- ^a Average of coefficients for 1929 and 1930, 1932 and 1933, 1933 and 1934.
- ^b Average of coefficients for 1929 and 1931, 1932 and 1934, 1934 and 1936.
- ^c Average of coefficients for 1932 and 1933, 1933 and 1934.
- ^d Average of coefficients for 1929 and 1930, 1932 and 1933, 1933 and 1934, 1934 and 1935, 1935 and 1936.
- ^e Average of coefficients for 1929 and 1931, 1932 and 1934, 1934 and 1936.
- ^f Average of coefficients for 33 different cities. The individual coefficients vary from .63 to .86, with 5 less than .7, 12 from .7 to .75, 12 from .75 to .8, and 4 from .8 to .86.
- ^g Average of six coefficients for all consecutive pairs of years from 1929 to 1935.
- ^h Average of coefficients for 1929 and 1931, 1933 and 1935.
- ⁱ Average of coefficients for 1929 and 1932, 1932 and 1935.
- ^j Average of coefficients for 1929 and 1933, 1931 and 1935.
- ^k Average of coefficients for 1929 and 1934, 1930 and 1935.
- ^l Average of six coefficients: for 1940 and 1941, 1941 and 1942 for three separate state groups: Illinois, Iowa, and Minnesota.
- ^m Average of three coefficients for separate state groups in Illinois, Iowa, and Minnesota.

Source:

Items 1 to 5

Milton Friedman and Simon Kuznets, *Income from Independent Professional Practice*, National Bureau of Economic Research, 1945, Table 56, p. 305.

Item 6

Horst Mendershausen, *Changes in Income Distribution during the Great Depression*, Studies in Income and Wealth, VII, National Bureau of Economic Research, 1946, Table 30, p. 90.

Item 7

Frank A. Hanna, Joseph A. Pechman, Sidney M. Lerner, *Analysis of Wisconsin Income*, Studies in Income and Wealth, IX, National Bureau of Economic Research, 1948, Part III, by Frank A. Hanna, "The Accounting Period and the Distribution of Income," Table 15, p. 232.

Item 8

Based on data from Survey of Consumer Finances reinterview sample collected by Michigan Survey Research Center for Federal Reserve Board. Coefficient taken from Margaret Reid, "The Relation of the Within-Group Transitory Component of Incomes to the Income Elasticity of Family Expenditures," unpublished paper.

Items 9 to 11

Ibid.

2. Empirical Evidence on P_y

A collection of computed correlation coefficients between incomes in different years is given in Table 18.⁹ These can be taken as estimates of the contribution of the permanent component in the corresponding years, under the variability assumption, or of the average contribution in the corresponding years, under the mean assumption. As expected, the coefficients decline with an increase in the number of years intervening between the years correlated. The decline is, however, on the whole moderate; the results are therefore not likely

⁹ The correlations summarized in Table 18 are all that, to the best of my knowledge, are available for the data in question. Omissions from the list—e.g. the absence of a correlation for physicians for 1930 and 1931—correspond with omissions in the original source.

EVIDENCE FROM INCOME DATA

to be greatly affected by the precise definition of the permanent component that is adopted.

The most striking feature of the table is the consistency among the correlation coefficients for such different groups and time periods, with the exception of farm families. The summary in Table 19 of the

TABLE 19
Summary of Correlation Coefficients in Table 18
for Three Principal Bodies of Nonfarm Data

<i>Number of Years Intervening between Years Correlated</i>	<i>Average Coefficient of Correlation for</i>		
	<i>Professions</i>	<i>Wisconsin Taxpayers</i>	<i>Urban Families</i>
0	.85	.83	.83
1	.80	.78	
2	.74	.76	
3		.71	.74
4		.70	
5		.69	

correlation coefficients for the three principal bodies of nonfarm data emphasizes this consistency.

The differences among the various groups are clearly small enough to be attributed to sampling variation. According to these data, for nonfarm families, the contribution of the permanent component to the variance of income cannot be set higher than about .85, on the broadest definition of the permanent component, nor lower than about .70 on a rather narrow definition. For a three year permanent component span—that is, for one year intervening between the years correlated—the relevant value of P_v is about .80. The correlation coefficients for farm families are distinctly smaller than for urban families, as general knowledge would lead one to expect. For the small and unrepresentative samples covered in Table 18—unfortunately the only ones for which we have data—the coefficient is between .4 and .5 for consecutive years and between .3 and .5 for nonconsecutive years with one year intervening.

Part of this difference between the coefficients for farm and nonfarm families may reflect the use of logarithms in computing the farm correlations and of absolute incomes in computing the others. Some bits of evidence suggest that for data like these, the computed correlation coefficient between the logarithms is generally lower than between the absolute values. But this can at most account for a small part of the difference.¹⁰

¹⁰ Margaret Reid has computed logarithmic correlations matching 14 of Mendershausen's 1929-33 correlations. The logarithmic correlation was lower in every case and the average correlation coefficient was .64 for the logarithms and .74 for the original values.

EVIDENCE FROM INCOME DATA

3. *Comparison of Estimates of P_y with Estimated Income Elasticity of Consumption*

On our hypothesis, the elasticity of consumption with respect to measured income computed from budget data is also equal to P_y , so the correlation coefficients in Tables 18 and 19 can be regarded as estimates of corresponding income elasticities.¹¹ The numerical values in these tables are clearly of the right order of magnitude to be estimates of the income elasticity of consumption. According to Table 1, the elasticity computed directly from the consumption-income data varies from .70 to .87 for various groups of nonfarm consumer units in the United States, with something like .83 as a reasonably typical value, and is .65 and .69 for two samples of farm units. For the nonfarm groups, these values are as good a summary of the correlations in Tables 18 and 19 as of the elasticities in Table 1. For the farm families, the correlations in Table 18 are lower than the elasticities in Table 3, but, as we shall see, so are elasticities computed for the particular farm groups covered by Table 18, so again the agreement is excellent.

The agreement between the values of P_y estimated from income data alone and from the regression of consumption expenditures on measured income must be regarded as strong evidence in favor of our hypothesis. Two considerations make this agreement particularly striking: first, most of the data underlying Table 18 are entirely independent of the data underlying the elasticities in Table 1; second, we were led to compare these two kinds of data as we have just done solely by our hypothesis and not by any previously noted similarity between the two magnitudes. To the best of my knowledge, no one has hitherto made a comparison of this kind, or indeed, of any kind, between these two kinds of data. It should perhaps be noted explicitly that there is nothing in the arithmetic of the computations to produce the observed measure of agreement. The correlation coefficients are constrained within the range -1 to $+1$; the elasticities can have any value from $-\infty$ to $+\infty$, though we know that empirically they are uniformly less than $+1$.

Our comparison has so far been without regard to the length of the horizon that defines the permanent component. We have been able to neglect this feature because, as noted, the effect of the length of the horizon on P_y is moderate, causing it to vary—for nonfarm groups—only from about .70 to about .85, so that leaving it undefined still gives a narrow enough range of values to provide an impressive

¹¹ This statement is exact for arithmetic linear regressions only at the mean point and for zero mean transitory components; for logarithmic linear regressions, it is exact more generally.

EVIDENCE FROM INCOME DATA

check on the coincidence predicted by our theory between the two sets of estimates of P_Y . Given this general coincidence, we can use these data to estimate the appropriate length of horizon. For nonfarm groups the correlation coefficients for consecutive years, which corresponds with a two-year horizon, are in general somewhat higher than the income elasticities computed from the budget data: one correlation coefficient is .67, the other six vary from .83 to .93 and the seven average .84. The elasticities for nonfarm groups vary from .70 to .87, the .70 being for 1944. The four others for the period after World War I are between .80 and .87. All eight average .80, the five for the period after World War I, .81. The correlation coefficients for nonconsecutive years with one year intervening, which corresponds with a three-year horizon, match the elasticities somewhat better: one is .63, the other five vary from .78 to .91, and the six average .80. On this evidence, a three-year horizon gives the definition of the permanent component that fits these data the best.¹²

Margaret Reid has made a more precise and detailed test of the relation between estimates of P_Y computed from income data alone (strictly speaking, P_Y , since she used logarithms throughout) and elasticities computed from consumption-income regressions.¹³ For a number of different groups of families for which income data were available for a number of consecutive years and expenditure data for one or more of these years, she has compared income elasticities computed from the budget data with estimates of P_Y computed from incomes in different years. Construction of both estimates for the same families eliminates one source of noncomparability that affected the preceding comparison between the elasticities in Table 1 and the correlation coefficients in Table 18. In addition, she has been able to estimate P_Y from the income data on the mean assumption and so get estimates for each year separately.

Much of Reid's data are for the farm families analyzed also by Tobin (see Chapter VI, section 4 above) and used in Chapter IV in considering the effect of change of income. As noted earlier, there is considerable doubt about the representativeness and accuracy of these data. Although these defects should affect the direct estimate of P_Y and the income elasticities in much the same way, and so do not

¹² One qualification that is required in connection with this comparison is that the elasticities were estimated from logarithmic regressions by graphic methods, the correlations computed from arithmetic data. It is not clear what the net effect of these differences is: the use of logarithms would probably reduce the correlations (see footnote 10); on the other hand, computation of the elasticities from all the data rather than graphic estimation would probably reduce the estimated elasticities (see Chapter IV, footnote 11). So these two differences probably offset one another, at least in part.

¹³ Described in Reid, "The Relation of the Within-Group Transitory Component of Income to the Income Elasticity of Family Expenditures."

EVIDENCE FROM INCOME DATA

destroy the value of the data for the present purpose, they probably introduce a good deal of variability into the results.

Figure 15, taken from Miss Reid's unpublished paper, summarizes the results. Both the income elasticities and the estimates of P_Y from income data were computed from the logarithms of the original observations. For such data, our hypothesis implies that the observed income elasticity of consumption should be equal to P_Y , provided, of course, that the "permanent component" is appropriately defined. If the data conformed precisely to this expectation, the observations would all fall on the diagonal lines in the panels.

In Panel 1, P_Y is estimated from data on income in two consecutive years, which implies the broadest definition of permanent component. The resulting values might be expected to be upper estimates of the appropriate P_Y or of the income elasticities. The results conform to this expectation. The cluster of points follows the pattern of the diagonal line, but tends to be to the right of it, an effect which overestimates of P_Y would produce. In judging this figure, it should be noted that (1) all but one of the points are for farm families, and the one exception is the point corresponding to the highest recorded value of P_Y , which is for the urban sample collected by the Michigan Survey Research Center; (2) the points are for years varying from 1937 to 1948; (3) the points for farm families are all computed from relatively small samples, varying from 60 to 229 families; (4) as just noted, at least some and perhaps many of the farm samples may be highly unrepresentative.

For some of the samples, data were available for three consecutive years. For these, Reid estimated P_Y from data for the first and third years, which implies a more restrictive definition of the permanent component. The results are plotted in Panel 2. As was to be expected, the substitution of a three-year for a two-year permanent component shifts the cluster of points to the left; one-third of the points are above the diagonal, whereas in Panel 1 only one-seventh are. The points in Panel 2 appear to follow the pattern of the straight line less well than in Panel 1; and taken at their face value, a flatter line than the diagonal seems called for. However, this difference should be given little weight. The appearance of flatness is produced entirely by points for the more dubious set of farm samples; the points for this set of samples alone show the same tendency in Panel 1, but it is there concealed by the larger number of other points in the chart.¹⁴

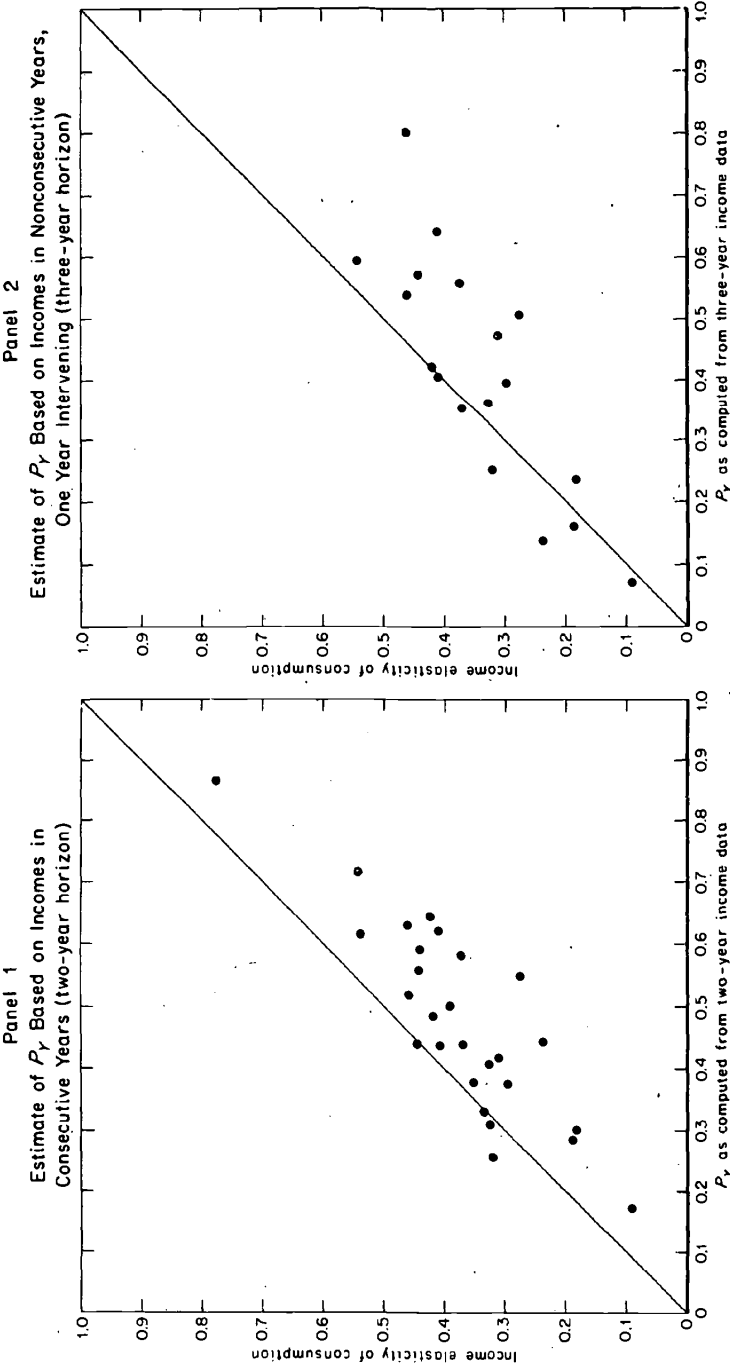
Panel 2 suggests that the appropriate definition of permanent component is for a period of three years or slightly longer. This is the

¹⁴ The points in question are all for the Farm Security Administration samples analyzed explicitly by Tobin and discussed above.

EVIDENCE FROM INCOME DATA

FIGURE 15

Relation between Computed Income Elasticity of Consumption and Computed Importance of Permanent Component of Income



Source: Margaret Reid, "The Relation of the Within-Group Transitory Component of Incomes to the Income Elasticity of Family Expenditures," unpublished paper.

EVIDENCE FROM INCOME DATA

same as the conclusion reached earlier from data, primarily for urban families, none of which is used in Panel 2. This common result is also consistent with the evidence from time series data. The expected-income regression computed in Chapter V between real per capita consumption and a weighted average of current and past incomes gave an estimate of 2.5 years for the average time lag or of 5 years for the effective horizon. This is longer than the horizon implied by the cross-section data. And it is plausible that it should be. Vagaries that reduce the effective horizon for individuals tend to cancel out in average data. It is encouraging to find such close agreement in the precise definition of permanent components suggested by three independent bodies of data.

After the preceding part of this section was written and circulated in mimeographed form, some additional data were transmitted to me by John Frechtling of the Federal Reserve Board on the relation between P_v as estimated from income data alone and the income elasticity of consumption as estimated from data on income and consumption (Table 20). These additional data are from a reinterview sample taken in connection with the 1953 Survey of Consumer Finances and cover the incomes and some wealth items of a limited number of spending units for 1951, 1952, and early 1953. Perhaps their most serious limitations for our purposes are: (1) Income is before rather than after taxes. (2) Data on all forms of savings are not available, so savings were approximated by the change in liquid assets plus the change in short-term debt. The major items omitted are savings in the form of real estate and other contractual items. (3) As is also the case with most of the other data used, no part of expenditures on durable consumer goods is treated as savings. What I have designated "consumption" in Table 20 is income less the indicated approximation to savings, which means that it includes personal taxes, some items of positive or negative savings, and expenditures on durable consumer goods.

The deficiencies in the definition of saving presumably explain why the average propensity in line 2 for independent business units is so much higher than in other data; the average propensities for the other two groups are not out of line. The income elasticities in line 4 seem somewhat higher than for most of the other studies we have examined, especially for the two nonfarm groups; this may well reflect not only the truncated saving definition but also the implicit inclusion of personal taxes as consumption. The alternative direct estimates of P_v in the last three lines of the table are strikingly similar to the income elasticities. The agreement is particularly close for the correlation coefficients in line 5, which are estimates of P_v under the variability

EVIDENCE FROM INCOME DATA

assumption. It is decidedly less close for the estimates of P_y under the mean assumption, particularly for 1952 in line 7, and at least as much weight should be given to this comparison, since some very limited evidence has led me to prefer slightly the mean assumption to the variability assumption.

While this piece of evidence agrees with the other evidence presented above in showing a close relation between the income elasticity as

TABLE 20

Alternative Estimates of P_y and Other Data from Survey of Consumer Finances, 1953 Reinterview Sample, for Three Occupational Groups

	<i>Independent Nonfarm Business</i>	<i>Farm Operators</i>	<i>Clerical and Sales</i>
1. Number of cases	83 ^a	99	132
Measures based on income and consumption data for 1952:			
2. Average propensity to consume	.99	.89	.95
3. Marginal propensity to consume ^b	.83	.62	.85
4. Income elasticity of consumption ^c	.83	.69	.90
Measures based on income data for 1951 and 1952:			
5. Correlation coefficient (estimate of P')	.83	.68	.88
6. Estimate ^d of P_{1951}	.85	.52	.94
7. Estimate ^d of P_{1952}	.81	.91	.83

^a Omits one case with exceptionally high income in 1951 and 1952.

^b Slope of computed arithmetic linear regression of consumption on income.

^c Marginal propensity divided by average propensity, i.e. elasticity at mean income.

^d Computed from formulas (7.1) in text.

Source:

Based on sums of observations, squares, and cross-products kindly made available to me by John Frechtling when he was with the Board of Governors of the Federal Reserve System. All computations are unweighted, which means that upper income groups are overrepresented.

computed from consumption-income data and estimates of P_y computed directly from data on income in two years, it differs in one important respect. The elasticities here are about the same as or larger than the correlation coefficients for successive years, whereas in the other comparisons they have been smaller. I am inclined to attribute this divergence to an overestimate of the elasticities as a result of both the truncated savings definition and the inclusion of personal taxes in consumption, but I have no independent evidence to test this conjecture.

EVIDENCE FROM INCOME DATA

4. *Correlation of the Ratio of Savings to Income in Consecutive Years*

In a discussion of the variability of consumer behavior based largely on the Survey of Consumer Finances reinterview sample for 1947 and 1948 (the source of the correlation coefficient in line 8 of Table 18), Katona gives a correlation table showing the relation between savings as a percentage of income in 1947 and in 1948 for 655 urban units.¹⁵ This table serves as an additional piece of evidence on the permanent income hypothesis, and, even more directly, on the appropriate length of horizon.

Let us treat these data as if they were for a group of consumer units that all had the same numerical value of k . Suppose transitory components of both consumption and income were zero for every consumer unit. The ratio of consumption to income (so of saving to income) would then, on our hypothesis, also be the same for all, namely k (or $1 - k$ for saving to income). For a given k , differences among consumer units in this ratio therefore reflect the effect of transitory components of income and consumption alone. If the relevant horizon were two years, factors would be regarded as transitory only if they affected income and consumption in just one year, so transitory components in successive years would be uncorrelated. It follows (as is demonstrated in the Appendix to this Chapter) that the ratio of consumption to income in the two years would also be uncorrelated.¹⁶ Suppose the horizon were three years. Some factors would then be regarded as transitory even though they affected income in two years, so transitory components in successive years would be correlated, though in years separated by a year they would be uncorrelated. The result would be to introduce correlation between the ratios of consumption to income in successive years, while leaving the ratios in two years separated by a year uncorrelated. In general, the longer the horizon, the higher tends to be the correlation between the ratios of consumption to income in consecutive years, and the longer the span of years for which a correlation exists.

It is clear, I trust, even from this brief sketch that, according to the permanent income hypothesis, the extent of correlation between incomes in pairs of years, both consecutive and nonconsecutive, and the length of the horizon are critical factors determining the extent of correlation between the ratios of consumption to income in two

¹⁵ George Katona, "Variability of Consumer Behavior and the Survey Method," Chapter II in Katona, Klein, Lansing, Morgan, *Contributions of Survey Methods to Economics*, p. 71.

¹⁶ Note that the correlation between the ratios of consumption to income in two years is numerically identical with the correlation between the corresponding ratios of saving to income.

EVIDENCE FROM INCOME DATA

consecutive years. It is perhaps less clear that such information alone is enough, on rather general assumptions, to determine the numerical value of the correlation coefficient. Yet the Appendix to this Chapter demonstrates that this is so. The correlation for Katona's table for a horizon of three years is there estimated in this way to be .25; the computed correlation is .26.¹⁷ We may perhaps regard this as a fourth piece of evidence for a three-year horizon.

Katona analyzes his table to determine whether particular patterns of savings ratios are systematically related to other characteristics of consumer units, whether and how the patterns carry over to particular forms of saving, and the like. Needless to say, his results are almost wholly negative, except where they are of the nature of arithmetical necessities.¹⁸ If our interpretation of his table is correct, his procedure is comparable with tossing a thousand fair coins twice, classifying the coins according as they come up heads both times, heads once and tails once, or tails both times; and then examining the resultant groups of coins to see why they behaved differently. The only respect in which this analogy is inexact is that there may be some systematic differences in k among the consumer units covered by the table. However, on the basis of our earlier quantitative results, such differences might be expected to be much smaller than the differences introduced by transitory components. It is as if the thousand coins, instead of all being perfectly fair, varied slightly in the probability of a head, or were mostly fair but included a few that were biased. It would take a good many more than two tosses apiece to classify the coins confidently by the probability of a head, or to isolate with confidence the few slightly biased coins.

Appendix:

Correlation between Savings Ratios in Two Consecutive Years

We shall deal throughout with the ratios of consumption to income rather than of savings to income. However, this amounts

¹⁷ Such close agreement is to be regarded as an accident. It does not, however, in any way reflect a choice of the assumptions of the calculations in the Appendix to fit the observed value; I estimated the correlation as .25 before I computed the actual correlation.

¹⁸ For example, he regards the table as revealing "one further significant fact. Repetitiveness, which is more frequent in positive than in negative saving, appears to be concentrated among people who save a small percentage of their income" (*ibid.*, p. 70). The greater frequency of repetitiveness in positive than in negative saving simply reflects the fact that in each year separately there are more positive than negative savers. Given that about two-thirds are positive savers and one-third zero or negative savers in each year, four-ninths would be positive savers in both years in the absence of any correlation, only one-ninth zero or negative savers in both. The explanation of the concentration of repetitiveness among small savers is the same: small savers are much the most numerous in both years; it would be phenomenal if they were not the most numerous among repeaters as well.

EVIDENCE FROM INCOME DATA

only to a change of coordinates and so does not affect the correlation.

Let us use the logarithmic variant of our hypothesis, and write out the equations in full, so that

$$(7.3) \quad \log c = \log c_p + \log c_t,$$

$$(7.4) \quad \log y = \log y_p + \log y_t.$$

Subtract (7.4) from (7.3) and express in arithmetic terms:

$$(7.5) \quad \frac{c}{y} = \frac{c_p c_t}{y_p y_t} = k \frac{c_t}{y_t}.$$

Suppose k is the same for all consumer units in question. The correlation between c/y in the two years and between c_t/y_t is then identical, since the only difference is in the unit of measure. So our problem reduces to the correlation between c_{t1}/y_{t1} and c_{t2}/y_{t2} , where the years are indicated in the subscripts. Expressed in logarithmic terms, our problem is to determine the correlation between $\log c_{t1} - \log y_{t1}$ and $\log c_{t2} - \log y_{t2}$, or between $C_{t1} - Y_{t1}$ and $C_{t2} - Y_{t2}$. For simplicity, assume the transitory components of both consumption and income to average zero in each year. Then

$$(7.6) \quad r_{(C_{t1}-Y_{t1})(C_{t2}-Y_{t2})} = \frac{E(C_{t1} - Y_{t1})(C_{t2} - Y_{t2})}{[E(C_{t1} - Y_{t1})^2 E(C_{t2} - Y_{t2})^2]^{\frac{1}{2}}}.$$

On our hypothesis,

$$(7.7) \quad r_{C_{t1}Y_{t1}} = r_{C_{t2}Y_{t2}} = r_{C_{t1}Y_{t2}} = r_{C_{t2}Y_{t1}} = 0.$$

For simplicity, assume further that

$$(7.8) \quad \begin{cases} \sigma_{C_{t1}}^2 = \sigma_{C_{t2}}^2 = \sigma_{C_t}^2, \\ \sigma_{Y_{t1}}^2 = \sigma_{Y_{t2}}^2 = \sigma_{Y_t}^2. \end{cases}$$

Expanding (7.6) and using (7.7) and (7.8), we get

$$(7.9) \quad r_{(C_{t1}-Y_{t1})(C_{t2}-Y_{t2})} = \frac{r_{C_{t1}C_{t2}}\sigma_{C_t}^2 + r_{Y_{t1}Y_{t2}}\sigma_{Y_t}^2}{\sigma_{C_t}^2 + \sigma_{Y_t}^2}.$$

Assume that

$$(7.10) \quad r_{C_{t1}C_{t2}} = r_{Y_{t1}Y_{t2}}$$

so that

$$(7.11) \quad r_{(C_{t1}-Y_{t1})(C_{t2}-Y_{t2})} = r_{Y_{t1}Y_{t2}}.$$

If the horizon were two years, transitory components of both consumption and income would be uncorrelated: indeed, this may be regarded as a definition of a two year horizon. In this case, (7.9) as well as the special case (7.11) would be zero.

EVIDENCE FROM INCOME DATA

The analysis in section 2 of the Appendix to Chapter 7 of *Income from Independent Professional Practice* can be used to give an estimate of $r_{Y_{t_1}Y_{t_2}}$. The term "transitory" was used there to designate a component affecting income in one year only, and "quasi-permanent", to designate components affecting income in more than one year. In the notation used there for a three-year horizon, our $Y_{t_1} = t'_1 + q'_{11} + q'_{12}$, our $Y_{t_2} = t'_2 + q'_{22} + q'_{23}$, where t'_1 and t'_2 are the components affecting year 1 and year 2 alone, q'_{11} and q'_{12} are two-year components affecting year 1 income and ending their effect in years 1 and 2 respectively, and q'_{22} and q'_{23} are two-year components affecting year 2 income and ending their effect in years 2 and 3 respectively (see *ibid.*, p. 353). If we extend (7.8) to apply to each component of Y_{t_1} and Y_{t_2} separately, we can make a transformation which will make $q'_{12} = q'_{22}$, since these are the quasi-permanent components produced by a factor common to both years; the other components are all uncorrelated between the two years. In consequence, the correlation between Y_{t_1} and Y_{t_2} is the variance of the common item over the common total variance, or

$$(7.12) \quad r_{Y_{t_1}Y_{t_2}} = \frac{\sigma_{q'_{22}}^2}{\sigma_{t'_1}^2 + \sigma_{q'_{23}}^2 + \sigma_{q'_{12}}^2}.$$

If we divide numerator and denominator by σ_Y^2 , and recall that our assumptions imply the variability assumption, this reduces to

$$(7.13) \quad r_{Y_{t_1}Y_{t_2}} = \frac{Q_{22}^*}{1 - P^*} = \frac{r_{12} - r_{13}}{1 - r_{13}},$$

where Q_{22}^* is the proportionate contribution of q_{22} , and P^* of the components lasting more than two years, estimated under the variability assumption; r_{12} is the correlation between incomes in years 1 and 2, that is, in two consecutive years; and r_{13} is the correlation between incomes in years 1 and 3, that is, in two nonconsecutive years, with one year intervening. From Tables 18 and 19, for urban units (Katona's table is for urban units), r_{12} appears to be approximately .85; r_{13} , approximately .80. Inserting in (7.13) we have

$$(7.14) \quad r_{Y_{t_1}Y_{t_2}} = \frac{.85 - .80}{1 - .80} = \frac{.5}{.20} = .25.$$

By (7.11), this is also an estimate of the correlation between ratios of savings in the two years.

This is a very rough approximation, not only because of the approximation involved in going from (7.9) to (7.11), but also because I have used correlations between arithmetic values to estimate logarithmic correlations.