Surjit S. Bhalla

Measurement Errors and the Permanent Income Hypothesis: Evidence from Rural India

Reprinted with permission from The American Economic Review, vol. 69, no. 3 (June 1970), pp. 295-307
Measurement Errors and the Permanent Income Hypothesis: Evidence From Rural India

By Surjit S. Bhalla*

The traditional, or Keynesian, consumption hypothesis postulated that (a) current income was a prime determinant of consumption; and (b) that the average propensity to consume declined with income. This hypothesis was rejected by Milton Friedman's permanent income hypothesis (PIH) which contended that (a) permanent income, and not current income, was the relevant determinant of consumption; and (b) that permanent consumption was proportional to permanent income (the proportionality hypothesis).

This paper attempts to test (and distinguish between) the two theories. Answers to two questions are sought: How important is the current income-permanent income distinction; and is the proportionality hypothesis, an important and controversial aspect of Friedman's theory, valid? The resolution of the former question has important implications for the understanding of habit and lags in consumption behavior and the efficacy of short-run macro-economic policies. The validity of the proportionality hypothesis has an important bearing on the recently popular controversy regarding the tradeoff between growth and equity in developing countries.2

The above two questions have been extensively investigated since the publication of Friedman's theory two decades ago.3 Nevertheless, with one exception, all these tests (including Friedman's) have been indirect and nonrigorous. The exception was an ingenious study by Nissan Liviatan, who used two-year panel Israeli data to test the PIH, and, in particular, the proportionality hypothesis.4 His study, however, suffered from a major drawback in that it failed to account for the bias caused by common measurement errors in consumption and income. This paper extends Liviatan's analysis by explicitly allowing for the presence of measurement errors in all the variables—income, consumption, and saving. The empirical estimation of the variances of these errors (source of bias) is possible due to the availability of independent estimates of consumption, savings, and income.5 The data base is a three-year panel survey, conducted by the National Council of Applied Economic Research (NCAER), New Delhi, of 4,118 households in rural India. (See Appendix A for description of data and definitions of variables.)

The incorporation of error variances in the therefore growth. (This is predicated on the assumption that growth in less developed countries (LDCs) is constrained by a lack of finance rather than of investment opportunities.) Thus, the PIH contends that there is no tradeoff between redistribution (equity) and growth.

2An excellent summary and interpretation of the voluminous literature on consumption behavior is contained in Thomas Mayer's book.

3Friedman accepted the compelling logic of Liviatan's tests and stated that "... if these [Liviatan's] results should be confirmed for other bodies of data, they would constitute relevant and significant evidence that the elasticity of permanent components is less than unity [the proportionality hypothesis]" (1963a, p. 63).

4Analogous conclusions are reached by the Modigliani-Brumberg life cycle hypothesis (LCH) of consumption behavior. The two theories, LCH and PIH, use different empirical methods but employ a common theoretical model of consumer behavior. The empirical techniques used in this paper apply directly to Friedman's formulation of the theory. Hence, the emphasis is on testing the PIH.

5These independent estimates allow for two measures of consumption—direct consumption and residual consumption (income minus savings). This "excess" of information is what makes the estimation of error variances possible. See Section II (and Appendix B) for details of methodology.
analysis, a unique aspect of this study, insures that unlike Liviatan's tests, the tests of the PIH in this paper are theoretically correct and most rigorous to date. Further, accounting for errors allows for the estimation of unbiased parameters of the Keynesian consumption function and makes possible a valid comparison with the PIH.

The plan of the paper is as follows. Section I presents a version of the PIH which is generalized to include the presence of pure measurement errors. Estimates of these error variances, as well as the uncorrected and corrected parameters of a Keynesian consumption function, are also presented. The “horizon,” a crucial parameter in the PIH, is defined, estimated, and corrected for measurement errors in Section II. A test of the proportionality hypothesis, and a comparison of the PIH with the standard Keynesian model, is presented in Section III.

I. A Generalized Version of the PIH

An important feature of this study is the distinction it makes between measurement errors proper, and the “measurement error” caused by the presence of transitory terms in the measured variables. Accounting for this distinction is necessary, and crucial, for a valid test of the PIH.

If measurement errors in income \( y \), consumption \( c \), and savings \( s \) are present, then a generalized version of the permanent income hypothesis would be

(1a) \( z^* = z' + z'' \)
(1b) \( z = z^* + z^0 \), \( z = y, c, s \)

where \( z^* \) represents the true value, \( z'(z'') \) represent the permanent (transitory) terms, \( z \) represents the measured value and \( z^0 \) the measurement error. Analogously, the variables can be expressed in logarithmic terms.\(^6\)

\(^6\)Lowercase letters represent natural numbers; upper case letters represent the logarithm (log) of a number. Equation (1) (arithmetic model) is implied by the assumption that transitory terms are additively distributed; if the variables are multiplicatively distributed (i.e., \( y = y'y'' \)) then a logarithmic model results. The former assumption results in a point estimate of the marginal propensity to consume; the latter in a point estimate of the consumption elasticity. Though both models are discussed, the development of the model is outlined in its logarithmic variant. Note that equation (2) does not contain an equation for savings, which can be negative for an individual household.

The assumption that the measurement errors are not correlated with each other needs to be qualified. Some items are common to consumption and income (for example, housing) whereas others are common to savings and income (nonmonetized investment). The magnitudes of these covariances are likely to be relatively small and are therefore ignored.

(2a) \( Z^* = Z' + Z'' \)
(2b) \( Z = Z^* + Z^0 \) \( Z = Y, C \)

The PIH assumes that the transitory components are independent of the permanent values, and each other, i.e.,

(3) \( \text{cov} (C', C'') = \text{cov} (Y', Y'') = \text{cov} (C'', Y'') = 0 \)

Analogous to (3), it may be assumed that measurement errors per se are not correlated with the true values, or with each other.\(^7\)

(4) \( \text{cov} (C^*, C^0) = \text{cov} (Y^*, Y^0) = \text{cov} (C^*, Y^0) = 0 \)

Apart from equations (2) and (3), a third relationship is postulated by the PIH; namely, that the permanent components are systematically related to each other, \( c^0 = ky' \), where \( k \) is assumed to be dependent on interest rates, tastes, composition of wealth, etc., but independent of \( y' \). This assertion implies a unitary elasticity between the permanent components, i.e.,

(5) \( C' = K + Y' \)

As is well known, estimation of the standard Keynesian function \( c = a + by \) or \( C = A + By \) in the absence of measurement errors yields a downwardly biased estimates of the permanent elasticity, \( \eta' \):

(6) \( \eta = \frac{\text{var } Y'}{\text{var } Y' + \text{var } Y''} < \eta' = 1 \)

Allowance for measurement errors in \( Y \) biases

\(^7\)The assumption that the measurement errors are not correlated with each other needs to be qualified. Some items are common to consumption and income (for example, housing) whereas others are common to savings and income (nonmonetized investment). The magnitudes of these covariances are likely to be relatively small and are therefore ignored.
downward the true measured elasticity $\eta^*$, as well:

$$\eta^* > \eta = \frac{\text{cov}(C, Y)}{(\text{var } Y^* + \text{var } Y^0)}$$

The relative magnitudes of $\text{var } Y^*$ and $\text{var } Y^0$ are unknown; lack of knowledge of the error variance $\text{var } Y^0$ can negate any comparison between the two consumption theories; that is, a comparison of the permanent elasticity $\eta^*$ with the traditional elasticity $\eta$. Further, a valid test of whether $\eta^*$ is equal to unity cannot be conducted. Fortunately, given certain assumptions, the magnitude of $\text{var } Y^0$ (and error variances in savings and consumption) can be estimated.

A maintained and general assumption in this paper is that measurement errors ($c^0, s^0$, and $y^0$) are independent of the true values, and have a zero mean. Thus, $E(z) = E(z^*)$ and $E(c) = E(y) - E(s)$, where $E(\cdot)$ is the expectation operator. Thus, the aggregate estimates of $c$ and its residual estimate, $c_r = y - s$ (alternatively $s$ and $s_r$), should be approximately equal. Such is not the case. The saving rates for the three years—1968–69, 1969–70, and 1970–71—are -0.6, 4.2, and 5.4 percent, respectively, for the direct measure, and 13.6, 14.6, and 16.3 percent, respectively, for the residual measure $s_r$.

Comparison with published national estimates also indicates that $s$ at the aggregate level is a more accurate measure of rural savings.\footnote{See the author for a detailed discussion of various savings estimates, and a comparison of $s$ and $s_r$.}

An equality between the two estimates of savings can be achieved $\text{ex post}$ by assuming that each household underestimates its consumption by a constant proportion $\alpha$. Thus, direct consumption $c$ is no longer equal to $c^* + c^0$, rather $c = \alpha c^* + c^0$, where $0 < \alpha < 1$. If $s$ is a "true" measure of aggregate savings, then $\alpha = E(c)/[E(y) - E(s)]$.

Given the $\text{ex post}$ equality between $c$ and $c_r$, and the value of $\alpha$, a "method of moments" approach can be used to derive the variances, $\text{var } y^0, \text{var } c^0$, and $\text{var } s^0$. (A key equation in the derivation is the identity between true values $y^* = c^* + s^*$.) Two different assumptions about the distribution of the errors are considered: (a) errors affect true values additively (arithmetic model, equation (1b)); and (b) errors affect true values proportionally (multiplicative model, equation (2b)). These two assumptions, though not exhaustive, cover a wide range of possibilities. Of the two, it is likely that the multiplicative assumption is more reasonable, that is, it is more likely that rich and poor households make the same proportional error, $X$ percent, rather than the same absolute error.

The solution of error variances is straightforward for the arithmetic model; for the more realistic multiplicative model, the methodology is much more involved. The difficulty arises primarily because the crucial identity $y^* = c^* + s^*$ does not exist in its (multiplicative) logarithmic transformation, i.e., $Y^* \neq C^* + S^*$. Nevertheless, an appropriate procedure exists. The solutions are outlined in Appendix B. The expressions for the income error variances for the two models are\footnote{Appendix B also contains expressions for errors in savings and consumption. Since the primary concern is with errors in the independent variable, income, only these are discussed in the rest of the paper.}

$$\text{(7a) } \text{var } y^0 = \text{var } y - \text{cov}(s, y) - \alpha^{-1} \text{cov}(c, y)$$

$$\text{(7b) } \text{var } Y^0 = \ln(1 + \text{var } M)$$

where $\text{var } M = \frac{\alpha \text{var } y - \text{cov}(c, y) - \alpha \text{cov}(s, y)}{\alpha [E(y)]^2 + \text{cov}(c, y) + \alpha \text{cov}(s, y)}$.
Table 1—Error Variances and Related Statistics

<table>
<thead>
<tr>
<th></th>
<th>Year 1 1968–69</th>
<th>Year 2 1969–70</th>
<th>Year 3 1970–71</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Income, Rs.</td>
<td>4544</td>
<td>4732</td>
<td>4988</td>
</tr>
<tr>
<td>Alpha (E(c)/[E(y) - E(s)])</td>
<td>.804</td>
<td>.850</td>
<td>.860</td>
</tr>
<tr>
<td>Direct Consumption, c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>.778</td>
<td>.759</td>
<td>.741</td>
</tr>
<tr>
<td>(MPC, \beta)</td>
<td>.39</td>
<td>.49</td>
<td>.47</td>
</tr>
<tr>
<td>(\beta) (adjusted by (\alpha))</td>
<td>.48</td>
<td>.58</td>
<td>.55</td>
</tr>
<tr>
<td>(\eta) (arithmetic model)</td>
<td>.50</td>
<td>.65</td>
<td>.63</td>
</tr>
<tr>
<td>(\eta) (log model)</td>
<td>.53</td>
<td>.71</td>
<td>.73</td>
</tr>
<tr>
<td>Residual Consumption, c,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>.97</td>
<td>.89</td>
<td>.86</td>
</tr>
<tr>
<td>(\beta)</td>
<td>.91</td>
<td>.70</td>
<td>.65</td>
</tr>
<tr>
<td>(\eta) (arithmetic model)</td>
<td>.94</td>
<td>.79</td>
<td>.76</td>
</tr>
<tr>
<td>(\eta) (log model)</td>
<td>.88</td>
<td>.80</td>
<td>.80</td>
</tr>
<tr>
<td>Errors and True (corrected) Values</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error Ratio ((\text{var } y^2/\text{var } y))</td>
<td>43.2</td>
<td>12.4</td>
<td>10.4</td>
</tr>
<tr>
<td>Typical Error (Rs.)</td>
<td>3014</td>
<td>1429</td>
<td>1430</td>
</tr>
<tr>
<td>Error Ratio ((\text{var } y^2/\text{var } y))</td>
<td>26.3</td>
<td>8.6</td>
<td>8.0</td>
</tr>
<tr>
<td>Typical Error (Percent)</td>
<td>49.6</td>
<td>23.2</td>
<td>21.7</td>
</tr>
<tr>
<td>True (\text{MPC, } \beta^*)</td>
<td>.84</td>
<td>.66</td>
<td>.61</td>
</tr>
<tr>
<td>True Elasticity, (\eta^*) (arithmetic model)</td>
<td>.87</td>
<td>.74</td>
<td>.71</td>
</tr>
<tr>
<td>True Elasticity, (\eta^*) (log model)</td>
<td>.72</td>
<td>.78</td>
<td>.80</td>
</tr>
</tbody>
</table>

Source: NCAER survey, cultivators only. For method of selection of observations, see Appendix A.

Equation (7) allows estimation of the error variances in terms of observed quantities; the estimates of these errors for the three years, and related data, are reported in Table 1. The results pertaining to the source of bias, the income error ratio, are striking. Errors (arithmetic model) account for as much as 43 percent of income variance in the first year of the survey; these decline significantly to 12 and 10 percent in the following two years. Even the more realistic assumption of proportional errors yields large errors for the first year—26 percent. Again, the error variance drops radically—9 and 8 percent in the second and third year, respectively.

A direct interpretation of measurement errors can be in terms of “typical” errors. Defining this to be one standard deviation, the errors are uncomfortably large: Rs. 3,014, Rs. 1,429, and Rs. 1,430 for the arithmetic model, and 50, 23, and 22 percent for the log model. Given that mean income is approximately Rs. 4,500, the error for the first year is especially discomforting.

The large error variance for the first year could have been caused by respondents being unfamiliar with survey procedures, suspicious of interviewers, etc. If these errors are typical of one-shot surveys in LDCs, then the analysis based on them is suspect.

There is no basis for judging these results since no prior estimates exist. The importance of the magnitude of even the low error variance for the third year, 8–10 percent, is emphasized by the fact that Friedman (1963a) contended that an error of 6 percent was sufficient to at least cast doubt upon, if not reverse, some of the results reached by Liviatan.

The measurement error variances are used to correct the observed coefficients and the results are also reported in Table 1. For the year 1970–71, the true marginal propensity to consume (\(\text{MPC} \beta^*\)) is 0.61 and the true elasticity (log model) \(\eta^*\) is .79. The observed \(\text{MPC}\) ranges between .55 and .65 (direct and residual estimates), and the observed elasticity between .73 and .80. The result pertaining to \(\eta^*\) is in close correspondence with other data and the Keynesian hypothesis, that is, the average propensity to consume (\(\text{APC}\)) of households declines with increases in current
income. If the PIH is mostly true, then the permanent elasticity should be significantly higher than \( \eta^* \) and (perhaps) equal to unity. But before \( \eta^* \) can be computed, it is necessary for the horizon to be estimated. This crucial parameter can also be biased by measurement errors; uncorrected and corrected estimates of the horizon are presented in the next section.

II. The Horizon

The proper length of the horizon has been a contentious issue in the literature on consumption theory. Part of the disagreement can be attributed to the fact that Friedman offers two definitions of the horizon without attempting to show any equivalence between the two. One definition is derived from considerations of wealth and its conversion into permanent income. Accordingly, if the subjective discount rate is \( r \), “the horizon can then be defined as \( 1/r \) or ‘the number of years purchase’ implied by the discount rate” (Friedman, 1963b, p.7). The second definition is directly related to the statistical formulation of the PIH and concerns itself with the separation of an income stream into its permanent and transitory components; “the length of time a factor must affect income before it is regarded as permanent is a measure of the length of the horizon” (Holbrook, p.750). According to the latter definition, a two-year horizon would imply that the transitory terms for the two years are not correlated. Analogously, “suppose the horizon were three years. Some factors would then be regarded as transitory even though they affected income in two years, so transitory terms in successive years would be correlated, though in years separated by a year they would be uncorrelated” (Friedman, 1957, p. 196).

A direct estimate of the second definition of the horizon can be achieved with NCAER data. Assumptions (additional to equations (1)-(5)) are necessary, but since these are identical to those made by Friedman, the tests are within the spirit (if not the letter) of the PIH. In particular, it is assumed that the permanent component of income changes (if at all) in a systematic manner from year to year. The systematic nature of change is either “that permanent components maintain the same ratio to the mean of the group in different years” (mean assumption), or “that the fraction of the total variability contributed by the permanent components is the same in successive years” (variability assumption), (Friedman, 1957, p. 184).

For the log model, mean assumption, the elasticity of incomes for any two years of the survey (indicated by the subscripts \( i \) and \( j \)), is

\[
(8) \quad \eta_{YiYj} = \frac{\text{cov}(Y_i, Y_j)}{\text{var} Y_j} = \frac{\text{cov}(Y'_i + Y''_i + Y'_0, Y'_j + Y''_j + Y'_0)}{\text{var} (Y'_j + Y''_j + Y'_0)}
\]

If \( Y'_i \) is systematically related to \( Y'_j \) (for example, \( Y'_i = \lambda Y'_j \) or \( Y'_i = \Gamma + Y'_j \)), then

\[
(8') \quad \eta_{YiYj} = \frac{\text{cov}(\Gamma + Y'_j, Y''_i + Y'_0, Y'_j + Y''_j + Y'_0)}{\text{var} Y_j} = \frac{(\text{var} Y'_j + \Delta Y'_i Y'_j + \Delta Y''_i Y'_j + \Delta Y'_0 Y'_j)}{\text{var} Y_j}
\]

where \( \Delta \)'s represent the respective covariances (i.e., \( \Delta Y'_i Y'_j = \text{cov}(Y'_i, Y'_j) \)). If these covariances are assumed to be zero, then (8') reduces to

\[
(9) \quad \eta_{YiYj} = \frac{\text{var} Y'_j}{\text{var} Y_j}
\]

Equation (9) is exactly as that obtained for the elasticity of consumption with respect to measured income, \( \eta \) (equation (6) ). Analogously, it can be shown that the correlation of incomes \( \rho_{YiYj} \) is the measure comparable to \( \eta \) under the variability assumption (see Friedman, 1957, p. 184):

\[
(9') \quad \rho_{YiYj} = \frac{\text{cov}(Y_i, Y_j)}{\{(\text{var} Y_i)^{\frac{1}{2}}(\text{var} Y_j)^{\frac{1}{2}}\}}
\]

This implies \( \eta = \frac{\text{cov}(C_i, Y_j)}{\text{var} Y_j} \). These “equivalences” can now be used to test for the length of the horizon. If transitory incomes
Table 2—Evidence on Consumer Horizon—Cultivator Households, Rural India

<table>
<thead>
<tr>
<th>Elasticities</th>
<th>Horizon</th>
<th>Correlations</th>
<th>Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Variability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assumption</td>
<td>Assumption</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arithmetic Model</th>
<th>Elasticiites</th>
<th>Horizon</th>
<th>Correlations</th>
<th>Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\eta$</td>
<td>$\eta_Y$</td>
<td>$\eta_{ry}$</td>
<td>$\rho_Y$</td>
</tr>
<tr>
<td>Direct Consumption</td>
<td>.63</td>
<td>.70</td>
<td>.69</td>
<td>H &gt; 3</td>
</tr>
<tr>
<td>Residual Consumption</td>
<td>.76</td>
<td>.70</td>
<td>.69</td>
<td>H &lt; 2</td>
</tr>
<tr>
<td>True Consumption</td>
<td>.71</td>
<td>.89</td>
<td>.57</td>
<td>2 &lt; H &lt; 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Log Model</th>
<th>Elasticiites</th>
<th>Horizon</th>
<th>Correlations</th>
<th>Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct Consumption</td>
<td>.73</td>
<td>.77</td>
<td>.83</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>Residual Consumption</td>
<td>.80</td>
<td>.77</td>
<td>.83</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>True Consumption</td>
<td>.794</td>
<td>$\gamma$</td>
<td>$\gamma$</td>
<td>$\beta$</td>
</tr>
</tbody>
</table>

*Coefficients move opposite to that "predicted." See fn. 12.

Correction of $\eta_{ry}$ is complex for log values if intertemporal correlation in measurement errors, equation (11), is allowed. The solution is not attempted.

are correlated, then for incomes measured within the horizon $\eta_{Y_Y} \neq \eta$ (a positive correlation implies $\eta_{Y_Y} > \eta$). Progressively larger distances between $i$ and $j$ should decrease the correlation towards zero. At the endpoints of the horizon, the correlation between transitory terms is equal to zero by definition, and $\eta_{Y_Y}$ (or $\rho_{Y_Y}$) equals $\eta$.

The discussion above is for transitory incomes proper. The presence of measurement errors can severely bias the above methodology and yield erroneous estimates of the horizon $H$. Equations (8') and (9') indicate that there are two error variances affecting the comparison: the measurement errors present in incomes for years $i$ and $j$; and the intertemporal covariance between the errors, $\text{cov}(Y^*, Y^*)$. Estimates for individual year errors $\text{var} Y_i^*$ and $\text{var} Y_i^*$ are available from Table 1. The intertemporal covariance $\text{cov}(Y_i^*, Y_i^*)$ may be considered insignificant since it involves correlations in errors a year, and two years, apart. However, it is possible that some errors are repeated and thus $\text{cov}(Y_i^*, Y_i^*) \neq 0$. An estimate of this covariance is possible if it is assumed that the intertemporal correlation of measurement errors is nonzero for the same variable and zero otherwise, i.e.,

$$ (10) \quad \text{cov}(w^0, z_i^0) \neq 0 \quad w = z; \\ \quad = 0 \quad w \neq z \\ w, z = c, s, i, j; \quad i, j = 1, 2, 3 $$

Incorporation of assumption (10) yields the following estimate for $\text{cov}(y_i^*, y_j^*)$:11

$$ (11) \quad \text{cov}(y_i^*, y_j^*) = \text{cov}(y_i, y_j) \\ \quad - \text{cov}(s, y_j) - \alpha^{-1} \text{cov}(c, y_j) $$

The measurement error corrected estimates of $H$ are presented with the uncorrected estimates in Table 2. These estimates are derived for the mean assumption (comparison of $\eta$ with $\eta_{Y_Y}$) and variability assumption (comparison of $\eta$ with $\rho_{Y_Y}$).12 It is seen that the "traditional," uncorrected (and biased) direct consumption estimate yields the result that the horizon (arithmetic model, mean assumption) is greater than three years. This result is in direct contradiction with most other estimates of the horizon. The importance of accounting for measurement errors is

11Derivation of equation (11) is contained in the author's working paper. The analogous solution for the log model, i.e., $\text{cov}(Y^*, Y^*)$ is more complex and is not attempted. The complexity arises from the fact that savings can be zero or negative, and as such cannot be (log) transformed.

12No estimates for the log model, mean assumption, are presented, since the elasticity two years apart $\eta_{Y_Y}$ is greater than the elasticity one year apart $\eta_{Y_Y}$. This "opposite" movement is easily explained if measurement errors in income are negatively correlated intertemporally. As mentioned in fn. 11, a test of the sign and magnitude of this covariance is complex and not attempted in this paper. Since estimates for the arithmetic model are available, this "error of omission" does not appear to be restrictive.
emphasized by the fact that correction for these errors reduces the estimate of $H$ to between two and three years. The variability assumption, log model, yields an estimate of $H$ equal to three years. All other estimates (corrected and uncorrected) place $H$ to be less than three years.

The "corrected" result that the upper bound for the horizon is three years is used in the next section to conduct strong tests of two components of the PIH: (a) how much larger is the permanent elasticity $\eta^*$ from the corrected elasticity $\eta$ (i.e., how important is the distinction between permanent and transitory income), and (b) is the proportionality hypothesis valid, i.e., is $\eta^* = 1$. The tests of the horizon itself were based on the assumption that $\eta^*$ is equal to unity. However, a "logical trap" is avoided if one notes that the method of estimating $H$—comparison of $\eta_{\text{corr}}$ and $\eta$—yields an upper bound of the horizon if the permanent consumption-income elasticity is less than unity. Since the bias is in the "right" direction, the rigorous tests outlined next are valid.

III. Estimates of the Permanent Income Model

A generalized expression for the measured consumption elasticity, with PIH assumptions, is

$$\eta = \frac{\text{var } Y'}{\text{var } Y' + \text{var } Y'' + \text{var } Y^0} < \eta^* < \eta' = 1$$

The fact that it is the presence of errors ($Y''$ and $Y^0$) in $Y$ which causes $\eta'$ to be downwardly biased suggests the use of an instrumental variable for $Y$. If this variable is represented by $X$, then the (instrument) consumption elasticity is

$$\eta^x = \frac{\text{cov}(C, X)}{\text{cov}(Y, X)}$$

If the proportionality hypothesis is valid ($C' = K + Y'$), then

$$\eta^x = \frac{(\text{cov}(Y', X') + \Delta C'' X'' + \Delta C^0 X^0)}{(\text{cov}(Y', X') + \Delta Y'' C'' + \Delta Y^0 X^0)}$$

where the $\Delta$'s represent the covariances between the transitory terms and measurement errors. And if the $\Delta$'s are zero, then $\eta^x$ reduces to the unitary permanent elasticity estimate $\eta'$.

Liviatan used the instrumental variable technique with two-year Israeli data and concluded that $\eta^x$ was less than unity. His instruments were the estimates of consumption and income of the "other" year. Unfortunately, his estimate of consumption was a residual estimate (i.e., $c_r = y - s$) and therefore common measurement errors were present in $c$ and $y$. Further, two-year data meant (if $H > 2$) that the $\Delta$'s were nonzero. Consequently, Liviatan's results were questioned by Friedman. Nevertheless, Liviatan's tests were accepted by Friedman, who stated that "it is most desirable that his [Liviatan's] analysis be applied to data not marred by common errors of measurement in income and consumption. It would further be highly desirable to have such data spanning at least a three-year period so that income and consumption for one year could be used as an instrumental variable for a year at least two years later or earlier" (1973b, p. 63).

Given the result that the upper bound for the horizon is three years, the three-year NCAER panel data fulfills all the requirements for a valid test. Further, the availability of separate estimates for consumption and saving suggests the presence of six proper instruments: income and two consumptions definitions for year 1 ($c$ and $c_r$) for each definition of consumption for year 3. If the PIH is correct, each instrument should yield an unbiased unitary permanent elasticity. This is only possible if all the $\Delta$'s are zero. The correlation amongst transitory terms is zero (by definition) for variables two years apart. The intertemporal measurement error correlations, however, may be nonzero. If so, then the use of first-year consumption $C_1$ (subscripts denote years) on a regression involving

The results of Table 3 are for a definition of consumption which excludes durable expenditures. This exclusion should minimize any possible correlation between the transitory components. All the tests of Table 3 were reestimated for a definition of consumption which included durable expenditure; virtually no difference was observed in the results.
Table 3—Consumption Function Parameters: All Households, 1970–71 (N = 2453)

<table>
<thead>
<tr>
<th></th>
<th>Arithmetic Model</th>
<th></th>
<th>Logarithmic Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Residual Direct</td>
<td></td>
<td>Residual Direct</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \beta ) ( \eta )</td>
<td></td>
<td>( \beta ) ( \eta )</td>
<td></td>
</tr>
<tr>
<td>Measured</td>
<td>.65 .76</td>
<td>.55 .62</td>
<td>.80 .73</td>
<td></td>
</tr>
<tr>
<td>( \text{Corrected} \ (\beta^<em>, \ \eta^</em>) )</td>
<td>(.007) (.010)</td>
<td>(.008) (.010)</td>
<td>(.008) (.008)</td>
<td></td>
</tr>
<tr>
<td>Permanen</td>
<td>.77† .89†</td>
<td>.73 .85</td>
<td>.89† .86</td>
<td></td>
</tr>
<tr>
<td>( \text{Instrument:} , c_i )</td>
<td>(.015) (.020)</td>
<td>(.017) (.020)</td>
<td>(.012) (.013)</td>
<td></td>
</tr>
<tr>
<td>Permanen</td>
<td>.70 .81</td>
<td>.66 .76</td>
<td>.85 .80</td>
<td></td>
</tr>
<tr>
<td>( \text{Instrument:} , c_i )</td>
<td>(.013) (.018)</td>
<td>(.015) (.018)</td>
<td>(.012) (.013)</td>
<td></td>
</tr>
<tr>
<td>Permanen</td>
<td>.71 .82</td>
<td>.65 .76</td>
<td>.84 .79</td>
<td></td>
</tr>
<tr>
<td>( \text{Instrument:} , y_i )</td>
<td>(.012) (.016)</td>
<td>(.014) (.016)</td>
<td>(.012) (.012)</td>
<td></td>
</tr>
</tbody>
</table>

Note: † refers to the "preferred" instrumental variable estimate. See text for details. Figures in parentheses are the standard errors of the coefficients. Arithmetic model: reported means and standard errors for \( \beta \) have been adjusted by \( a \); reported elasticities \( \eta \) are computed at the mean value for consumption. The standard errors are valid under the assumption that the \( \text{APC} \) is known, and is equal to a constant, \( k \).

\( C_i \) and \( Y_i \), for instance, would imply that \( \eta^x \) would be biased due to the presence of \( \Delta C_i^3 C_i^3 \); similarly, if \( Y \) is used as an instrument, i.e., \( \Delta Y_i^3 Y_i^3 \neq 0 \). However, if the assumption of equation (10) is accepted, then an instrument does exist whose use would yield an unbiased estimate of the permanent elasticity, \( \eta^p \) (i.e., \( \Delta S = 0 \)). This instrument is direct consumption, year 1, on a regression between residual consumption and income, year 3, i.e., \( \eta^p \) where

\[ \eta^p = \text{cov}(C_i, C_i) / \text{cov}(Y_3, C_i) \]

In this instance, only nonsimilar variables are involved and the correlation amongst these errors is assumed to be zero. Thus,

\[ \eta^p = (\text{cov}(Y_i^3, C_i^3) + \Delta C_i^3 C_i^3 + \Delta C_i^3 C_i^3) / (\text{cov}(Y_i^3, C_i^3) + \Delta Y_i^3 C_i^3 + \Delta Y_i^3 C_i^3) \]

\[ = \text{cov}(Y_i^3, C_i^3)/\text{cov}(Y_i^3, C_i^3) = 1 \]

Table 3 presents the results for both the arithmetic and log model of consumer behavior, and for all six instruments for income year 3. The two models yield point estimates of the permanent marginal propensity to consume, \( \beta^* \), and the permanent consumption elasticity, \( \eta^p \). Estimates of the true measured propensity and elasticity (from Table 1) are also presented. The results of Table 3 strongly reject the proportionality hypothesis. All equations, for both models, yield an elasticity which is considerably less than unity. However, the PIH estimates do tend to be larger than the corresponding Keynesian, current income estimates. The preferred estimate \( \eta^p \) is 25 percent larger, arithmetic model, and 13 percent larger, log model, than the corresponding traditional or Keynesian elasticity, \( \eta^* \). Thus, the general conclusion that emerges is that Friedman's contention of a "lag" in consumption behavior is justified, but not the assertion that permanent consumption is proportional to permanent income.

An aggregative relationship as estimated in Table 3 may not be valid. This is due to the presumed presence of "subsistence" households in rural India. These households are definitionally constrained to consume their entire income, i.e., their \( \text{APC} = k = 1 \). Households beyond the subsistence level presumably have an \( \text{APC} < 1 \). An explicit

---

14The estimation of \( \eta^p \) for the arithmetic model is computed under the assumption that the average propensity to consume \( k \) is a known constant. Thus, \( \eta^p = \beta^p/k = \beta^p/k \).

15All tests reported in Tables 3 and 4 were conducted in reverse, i.e., third-year variables as instruments for the first-year relationship. The results were qualitatively identical to those reported in this paper. These results further strengthen the conclusions of this paper.

16Arnold Zellner discusses this "weakness" or nongenerality of the permanent income hypothesis.
dependence between the APC and the permanent income level is therefore built into the model, and an elasticity less than one is the expected result. Thus, a test of the proportionality hypothesis is inappropriate.

Identification of a subsistence level, however, is difficult. The subject has been discussed at length in the Indian literature and the consensus seems to be that an annual income of Rs. 450 per capita, 1970-71 prices (corresponding to Rs. 15-20 per month, 1960-61 prices) adequately describes the subsistence level. If it is assumed that three-year average per capita income $y_a$ adequately reflects permanent status, then $y_a \leq Rs. 500$ should conservatively separate households into subsistence-nonsubsistence categories. If consumption functions are estimated for each group separately, then the prior expectation would be that (a) $\eta$ for subsistence households is unity (they are constrained to consume their entire income) and (b) $\eta$ for nonsubsistence households is also unity, if the PIH is valid.

Table 4 contains the results for the two classifications. For subsistence level households, the unbiased permanent elasticity is very close to one, and never significantly different from one. The most interesting result regarding these households is not that the elasticity is equal to one, but rather that the permanent marginal propensity to consume income is not different from one. These results conform very well to any prior expectation about subsistence level behavior. Subsistence households, if properly defined, should have a marginal propensity equal to one. Indeed, my results suggest an alternative definition of a subsistence level; in particular, that such a level (or range) is one after which the permanent marginal propensity becomes less than one.

The results for nonsubsistence households represent the strongest possible test of the PIH. The a priori dependence between the propensity to consume and permanent income has been "purged" from the data. The results, however, are not favorable to the proportionality assumption. For all instruments (arithmetic and logarithmic models), the elasticities are significantly less than one at the 1 percent level of confidence. These elasticities
range from .87 to .92, log model. The preferred result \( \eta^* \) yields \( \eta' \) equal to .92 which is significantly less than unity, but is also considerably larger than the corrected Keynesian elasticity, \( \eta^* = .79 \).

IV. Conclusions

This paper extends Friedman's permanent income model by explicitly allowing for the distinction between pure measurement errors and transitory terms in the observed variables. Incorporation of this distinction in the theoretical (and empirical) framework is necessary for valid, and direct, tests of the permanent income hypothesis. These tests do not need the measurement error associated with each individual's income per se; rather, only the variance of these errors is necessary.

This extended model not only allows for proper tests of the PIH (i.e., those that do not incorporate assumptions additional to the PIH) but also makes possible a correct comparison between the traditional and permanent income theory of consumer behavior. Neither hypothesis is supported in toto. The contention of the PIH that measured consumption elasticities are downwardly biased estimates of the true permanent elasticities is supported by the data—the difference between the two is large and in a direction predicted by the PIH, i.e., \( \eta' > \eta^* \).

However, a major and controversial aspect of the PIH is strongly rejected by the data—the elasticity between permanent components is less than unity. This result is subject to all the usual reservations about the applicability of a theory developed in the West for the households of rural India. However, if the permanent income hypothesis is general in its construction (and it is), then these results constitute one of the few, relatively unambiguous, refutations of the proposition that consumption is proportional to permanent income.

Separate consumption functions were estimated for subsistence and nonsubsistence households. The latter set of households were found to have not only a lower consumption elasticity but also a lower permanent marginal propensity to consume. It is an oft-noted result that the APC and MPC decline with increases in current income. The observed decline of MPC out of permanent income suggests that income redistribution policies are likely to have, ceteris paribus, a negative impact on the supply of household savings, and consequently growth, in the LDCs.

The importance of the bias caused by measurement errors is revealed at all stages of the analysis. This is not surprising given the magnitude of the error variances that are observed. These variances are found to be 43 and 26 percent of income variance (arithmetic and multiplicative errors, respectively) for the first year of the survey. This rather large magnitude suggests that researchers should be more cautious than usual with one-shot survey data, especially if these data are collected in rural areas. The error variances are observed to drop radically (to about 8–10 percent) for the second and third years of the survey, thereby suggesting that there is a "learning by doing" aspect to the collection and responses of survey data.

As a by-product, the results of this paper indicate that the length of the consumer's horizon is closer to Friedman's three-year estimate than the short horizons observed by Holbrook. The horizon was found to be between two and three years, but certainly not greater than three.

APPENDIX A: DATA AND DEFINITIONS

Data. The data are based on a panel survey of 4,118 households in rural India, 1968–69 to 1970–71. The survey, known as the Additional Rural Income Survey, was conducted by the National Council for Applied Economic Research (NCAER), New Delhi. The survey over-sampled high-income households and gathered data primarily on the pattern of income, consumption and savings of rural households.

For purposes of analysis, only households that were cultivators (self-cultivation on owned or leased land greater than .05 acres) for all three years of the survey were selected. This reduced the sample size from 4,118 to 2,532. Further, households with negative incomes and/or savings (change in net worth) that were estimated to be greater than income
for any year of the survey were excluded from analysis. This elimination reduces the sample to 2,453 households.

Definitions:
(a) Income \( y \)—The income of a household is defined as the total of the earnings of all the members of a household during a reference period. This income can be business income (farm or otherwise), wages, rents (land and house property), interest and dividends on financial investments, and pension and regular contributions.

(b) Savings \( s \)—The savings of a household is defined as the change in its net worth. This figure is adjusted for capital transfers. In other words, household savings \( s \) is defined to be \( s = dA - dL - dK \), where \( dA \) = gross change in the value of physical and financial assets, \( dL \) = net change in liabilities, and \( dK \) = net inflow of capital transfers. Consumer durables and nonmonetized investment are included in \( dA \). Savings in the form of currency or gold and silver are not included due to lack of reliable data; nor has any adjustment been made for capital gains or losses incurred by the household. Depreciation on assets is also ignored.

(c) Consumption \( c \)—In addition to the data on income and savings, the NCAER survey collected independent information on the consumption of households. About twenty-five food items, ten nonfood items (fuel, clothing, medicines, etc.), and "other" items (marriages, funerals, unexpected travel, etc.) comprise the information on consumption expenditures. Addition of these items yields one estimate of consumption \( c \). Subtraction of savings from income yields a residual estimate of \( c, c = y - s \).

APPENDIX B: DERIVATION OF MEASUREMENT ERROR VARIANCES

Measurement errors are presumed to exist in all three variables: consumption, savings, and income. These errors are assumed to be related either additively to the true values (arithmetic model) or multiplicatively (logarithmic model). A "method of moments" approach is used to derive the respective error variances; the notation is identical to that used in the text. The reader is referred to the author’s working paper for a more detailed derivation.

Arithmetic Model: Errors are assumed to be distributed independently of the true values, of each other, and to have zero mean. Thus,

\[
\begin{align*}
(A1) & \quad z = z^* + z^0 \quad z = c, s, y \\
(A2) & \quad \text{cov}(z^*, z^0) = 0 \\
(A3) & \quad E(z^0) = 0; E(z) = E(z^*)
\end{align*}
\]

As discussed in the text, two kinds of errors are present in consumption; a systematic underestimation by a factor \( \alpha \), and a random error \( c^0 \). Consequently (A1) is modified for consumption as

\[
(A1') \quad c = \alpha c^* + c^0, \quad 0 < \alpha < 1
\]

The real values are related by the identity

\[
(A4) \quad y^* = c^* + s^*
\]

Given assumptions (A1) and (A2), the observed variances and covariances are

\[
\begin{align*}
(A5a) & \quad \text{var} z = \text{var} z^* + \text{var} z^0 \\
& \quad z = y, s \\
(A5b) & \quad \text{var} c = \alpha^2 \text{var} c^* + \text{var} c^0 \\
(A5c) & \quad \text{cov}(s, y) = \text{cov}(s^* + s^0, y^* + y^0) \\
& \quad = \text{cov}(s^*, y^*) \\
(A5d) & \quad \text{cov}(c, y) = \text{cov}(\alpha c^* + c^0, y^* + y^0) \\
& \quad = \alpha \text{cov}(c^*, y^*)
\end{align*}
\]

(Strictly speaking, probability limits should be used for representing observed variables in (A5). For notational convenience, this usage is suppressed.)

Equations (A1)–(A5) and particularly the use of the identity \( y^* = c^* + s^* \) are sufficient to isolate the error variances. In terms of observed values, they are

\[
\begin{align*}
(A6a) & \quad \text{var} y^0 = \text{var} y - \text{cov}(s, y) \\
& \quad - \alpha^{-1} \text{cov}(c, y) \\
(A6b) & \quad \text{var} s^0 = \text{var} s + \alpha^{-1} \text{cov}(c, s) \\
& \quad - \text{cov}(s, y) \\
(A6c) & \quad \text{var} c^0 = \text{var} c + \alpha \text{cov}(c, s) \\
& \quad - \alpha \text{cov}(c, y)
\end{align*}
\]
Logarithmic Model: If measurement errors are proportional, rather than additive, the equations (A1)–(A3) are replaced by

\[ z = z^* z^0 \quad z = y, s \]
\[ c = ac^* + c^0 \]
\[ \text{cov}(z^*, z^0) = 0 \quad z = y, s, c \]
\[ E(z^0) = 1; \quad E(z) = E(z^*) \]

If incomes are lognormally distributed, then a log transformation of the above equations is preferable. (Equation (A12) does not follow from assumption (A9). The problems posed by this nonequivalence are minor and dealt with later.)

\[ Z = Z^* + Z^0 \quad Z = Y, S \]
\[ C = A + C^* + C^0 \quad (A = \ln \alpha) \]
\[ \text{cov}(Z^*, Z^0) = 0 \]
\[ E(Z^0) = 0 \]

Estimation of \( \text{var} Z^0 \), etc. is desirable to remove the bias from regressions involving the log transformations of variables (for example, equation (6)). Though equations (A10)–(A12) are identical to equations (A1)–(A3), a parallel, straightforward solution of \( \text{var} Z^0 \) does not exist. This is due to the fact that the identity \( y^* = c^* + s^* \) does not exist in the log transformation, i.e., \( Y^* \neq C^* + S^* \). An indirect procedure is employed instead. The parameters of the distribution of \( Z^0 \) are derived first; the log transformation of these distributions yield estimates of \( \text{var} Z^0 \).

The multiplicative errors of equation (A7) are assumed to be distributed independently of the true values. According to a theorem by Leo Goodman,

\[ \text{var} z = [E(z^0)]^2 \text{var} z^* + [E(z^*)]^2 \text{var} z^0 + \text{var} z^* \text{var} z^0 \]

Equations (A4), (A5c), and (A5d) also hold for the multiplicative error assumption. The system of equations can now be solved for \( \text{var} z^0 \), and results in

\[ \text{var} y^0 = \frac{\alpha \text{var} y - \text{cov}(c, y) - \alpha \text{cov}(s, y)}{\alpha[E(y)]^2 + \text{cov}(c, y) + \alpha \text{cov}(s, y)} \]

These variables are derived under the assumption that errors \( z^0 \) are distributed with mean one, rather than lognormally distributed with mean zero. The established relationship between normal and lognormal distributions (see J. Aitchison and J.A.C. Brown) allows the derivation of \( \text{var} z^0 \).

\[ \sigma^2 = \text{exp} (\mu + 0.5\sigma^2) \]

\[ \sigma^2 = \text{exp} (2\mu + \sigma^2) (\text{exp} (\sigma^2) - 1) \]

This relationship, the assumption \( E(z^0) = 1 \), and some algebra yields expressions for \( \sigma^2 \) or \( \text{var} Z^0 \):

\[ \sigma^2 = \text{ln} (1 + \sigma^2) \]

where \( \sigma^2 = \text{var} z^0 \), and expressions for \( \text{var} z^0 \), in terms of observed parameters, are as given in equations (A14)–(A16).

REFERENCES


The full range of World Bank publications, both free and for sale, is described in the World Bank Catalog of Publications, and of the continuing research program of the World Bank, in World Bank Research Program: Abstracts of Current Studies. The most recent edition of each is available without charge from:

PUBLICATIONS UNIT
THE WORLD BANK
1818 H STREET, N.W.
WASHINGTON, D.C. 20433
U.S.A.