BANCA D'ITALIA

Temi di discussione

del Servizio Studi

The Microeconomics and Macroeconomics of the Permanent Income Hypothesis

by Angus Deaton

Numero 166 - Marzo 1992

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

The Microeconomics and Macroeconomics of the Permanent Income Hypothesis

by Angus Deaton

Numero 166 - Marzo 1992

The Microeconomics and Macroeconomics of the Permanent Income Hypothesis

by Angus Deaton (*)

Abstract

Much of the literature on consumption is concerned with the interpretation of the aggregate data directly in terms of the permanent income hypothesis as derived for a single individual or household. So considered, the permanent income hypothesis is typically rejected. In this paper, I look at the microeconomic evidence, to see whether the theory fares better there, and whether we might not be able to make progress in understanding the macroeconimics of consumption and saving by working from the bottom up. The literature to date has typically found it much more difficult to reject the permanent income hypothesis on micro data, and those tests that have encountered negative results can typically be "explained" by econometric or data problems. Furthermore, it is possible to construct realistic models where the permanent income hypothesis is true for each agent, but where the model appears to fail in the aggregate, much as it does in the actual data.

(*) Princeton University, Department of Economics.
This paper was presented at the Workshop on "Savings in Italy: Past and Future Trends, Household and Government Behaviour", organized by the Bank of Italy and held in Rome on 16-17 January 1992. It is a self-contained version of Chapter 5, "Macroeconomics and Microeconomics" of the forthcoming book Understanding Consumption.

Much of the literature on consumption, arguably too much, is concerned with the interpretation of the aggregate data directly in terms of the permanent income hypothesis as derived for a single individual or household. So considered, the permanent income hypothesis is typically rejected. Consumption is "excessively sensitive" to current income. More precisely, changes in consumption are positively correlated with previously predictable changes in current income, something that should not happen if the representative agent looks forward using all the information that is currently available. This result, first documented by Flavin (1981), has been widely replicated since, and appears to hold, not only for aggregate time series data from the United States, but also for most other countries, see particularly Campbell and Mankiw (1991).

In this paper I look at the microeconomic evidence, to see whether the theory fares better there, and whether we might not be able to make progress in understanding the macroeconomics of consumption and saving by working from the bottom up. In Section 1, I look at a range of studies that have used microeconomic data to look at the behavior of consumption and income. These studies have not produced the sort of widely agreed stylized facts that have come out of the macroeconomic literature. Some authors find no conflict with the theory, while others have interpreted their findings as reinforcing at the household level the rejections of the theory in the aggregate data. Certainly, and at least as far as short-run behavior is concerned, the formal econometric analysis of the micro data does not appear to generate the obvious inconsistencies with the theory that are encountered in the macroeconomic data. Even so, there is a good deal of less formal evidence that liquidity constraints are important. But given the greater ambiguity of the micro data, it is worth exploring whether the macroeconomic problems are not generated somehow in the aggregation, by a failure of one of the assumptions required for the micro results to go through to the macro level. The crucial assumptions required for aggregation of the permanent income hypothesis are that people (or at least households) live for ever, and that aggregate variables are known to the individual agents. Relaxing either one of these assumptions generates macro models that differ in interesting ways from their micro parents, and that go at least some way towards explaining the aggregate findings, even if the theory holds good at the individual of household level. The two aggregation issues, finite lives and heterogenous information, are discussed in Section 2.

Although it seems likely that credit constraints are an important feature of reality for many households, the formal econometric tests on the microeconomic data do not provide overwhelming evidence of the effects of liquidity constraints on consumption. There are several findings that can be (and are) interpreted as revealing an inability to borrow, but most can also be interpreted in ways that do not threaten simple models of optimal intertemporal allocation. The aggregation results are also consistent with a story in which individuals obey the permanent income hypothesis, but in which aggregation "problems" cause average behavior not to conform to that of a representative permanent income consumer. Representative agents live too long and know too much, and the recognition of finite lives and heterogeneous and limited information makes it unlikely that the truth of the permanent income theory would be revealed in the aggregate data. There is therefore little in this review that would change the views of a believer in the theory. There is nevertheless a great deal of informal evidence that credit constraints are important, some of it briefly reviewed below, and the formal tests are as consistent with the presence of credit constraints as they are with their absence. A test that sharply discriminates between the two explanations will probably have to wait until we have longer panel data than those that are currently available.

1 Evidence from the microeconomic data

General problems of data and methodology

Microeconomic studies have the immediately appeal that they use the data that are appropriate for the theory. The theoretical framework of the intertemporal choice applies to an individual or a household and not to an aggregate or average of all individuals or households in the United States or elsewhere. Nevertheless, there are immediate practical problems, problems that are different from those encountered with the time-series data, but real nonetheless. One of the virtues of a representative agent is that the process of aggregation has destroyed all individual personality. The representative agent is neither young nor old, is certainly neither male nor female, has a uniform and roughly constant number of perpetually youthful children, all of which characteristics can be more or less ignored in the estimation. In the micro data, by contrast, the econometrician is immediately confronted with the differences between the individuals in the sample, and with the need to control for these differences if any progress if to be made in understanding their consumption. The convenient fiction of an invariant felicity function, with age doing no more than adding a discount factor, may be convenient for aggregate data, but it is nonsense at the micro level. Age and family composition matter, as do a host of other possible variables such as race, education, place of residence, and occupation. Indeed, diversity is so obviously important that it is hard to justify models that do not allow for the presence of unobservable individual fixed effects, effects that are certainly correlated with the income and consumption variables that concern us, and the introduction of which generates sometimes intractable problems of statistical inference.

Aggregation not only smooths away individual idiosyncracies, it also eliminates or much reduces the effects of measurement error. No one who has looked at the year to year variation in reported consumption and income in a micro data set comes away without being convinced that much of the variation is measurement error, see in particular Altonji and Siow's (1987) description and analysis of the income and food consumption data from the Panel Study of Income Dynamics (PSID). Income changes in the micro economic data typically display negative autocorrelation, both over years, MaCurdy (1982) and Abowd and Card (1985), and over quarters, Pischke (1991). This could reflect the importance of transitory incomes, or it could reflect the importance of random measurement error; they are not easily told apart. Furthermore, the aggregate quarterly income changes are *positively* autocorrelated over time, and the difference needs to be part of any complete story of micro and macro consumption and income.

It is also the case that many household surveys of income and expenditure do not generate plausible figures for saving. Most likely, the problems are again to do with measurement error; small percentage errors in either or both of the two large magnitudes, income or consumption, will cause much larger proportionate errors in measures of the small difference between them. Moreover, survey data often show households, particularly poor households, spending more than they earn. For many LDC's, household surveys record the bottom 50–80% of the income distribution as dissaving, and the phenomenon is so widespread that it is hard to attribute it to the surveys having been collected only in years of abnormally low income, see for example Visaria (1980). Such an effect is predicted by random measurement error; the slope of the regression function of consumption on mismeasured income is biased towards downward, generating apparent dissaving at low incomes, and apparent saving at high incomes. This is of course exactly the same effect that Friedman's (1957) permanent income theory uses to explain the positive intercept and slope less than unity in the cross-section 'Keynesian' regression of consumption on income. Even so, there is also a suspicion that there are systematic downward biases in the measurement of income, especially among those who are self-employed, and for whom it is extremely difficult even to define income, let alone to measure it.

Finally, and perhaps most seriously, there is a real lack of household survey data that are suitable for testing the predictions of consumption theory. The ideal would be long time-series data that track individual households. The PSID has been recording income data for a panel of households since 1967, but it collects data neither on saving nor on total consumption. Data are available from the PSID for consumption of food, currently some 17% of total consumption in the National Accounts, and, as we shall see, several authors have used these data to model aggregate consumption, typically under the assumption that the elasticity of food consumption (or the marginal propensity to spend on food) is a constant. The Consumer Expenditure Survey conducted by the United States Bureau of Labor Statistics is not a panel, but households are visited on five occasions over a fifteen month period, and are asked about income and consumption on two of these visits. This type of situation is perhaps the most common, with either a single cross-section containing some element of panel data, or a short panel, but in either case there will be few observations per household, where in many cases few means two. In such cases, the time-series variation that would ideally identify the model has to be replaced by cross-sectional variation, and this cannot be done without additional and (as we shall see) dangerous assumptions.

One other possibility should be noted. In some cases, there are household income and expenditure surveys that have been in the field for many years, so that, although individual households cannot be tracked, we have a time series of independent cross-sections. The Family Expenditure Survey in the United Kingdom is perhaps the most notable example, and has been collecting annual data on some 7000 households a year since 1954. With this kind of data, it is possible to construct synthetic cohorts following the method introduced in Browning, Deaton, and Irish (1985) and Deaton (1985). Although no household or group of households is observed more than once, we can think of a constant population of households, from which the successive surveys are drawing different samples. For example, think of the cohort of men born in 1945. In the 1975 sample, there is a subsample of all male 30 year olds, and averages can be calculated of their income, consumption, hours worked, or whatever is the variable of interest. In the 1976 survey, the procedure is repeated for 31 year olds, in 1977 for 32 year olds, and so on. In this way, we can track a cohort through their sample averages. For some purposes, and if the subsamples are large, the sample averages may be precise enough to be analyzed as if they were panel data. Otherwise, the sampling errors can be explicitly taken into account using an appropriate errors in variables estimator. Although this procedure uses (semi-) aggregated data, there is no aggregation problem, because averages can be computed for whatever function of the data is desired, for logarithms as well as levels, or whatever the functional form may be. But note again that, as with short panel data, we are immediately embroiled in econometric complexities because we do not have the data, long panel data, that would be ideal for testing the theory of consumption.

2F

Hall and Mishkin's study

One of the most frequently cited and influential of the papers on the microeconomic data is a study of food expenditure in the PSID by Hall and Mishkin (1982). The general approach is closely related to that used by Flavin (1981), with due adaptation for the nature of the data. In Flavin's study, consumption and income are detrended prior to estimation on the assumption that income possesses a deterministic trend. In Hall and Mishkin's paper, the individual components in income and consumption are allowed for by running preliminary regressions on household characteristics, and then working with the residuals from the two regressions. The change in income is assumed to have a deterministic component that is a function of age, age squared, time, and the changes in the numbers of adults and children in the household. The corresponding equation for the change in consumption includes these variables plus a relative price term. The income regression is designed to identify that part of permanent income which can be calculated by the household in advance, so that the change in the consumption residual should be determined only by innovations in the income residual. These issues dealt with, it is possible to focus on the topic of interest, which is the short-run dynamic response of consumption to income.

Write $\tilde{y}_t$ and $\tilde{c}_t$ for labor income and consumption with the individual effects swept out. Hall and Mishkin then propose the following time series model for income:

$$\tilde{y}_t = y_t^L + y_t^S \tag{1}$$

where the superscripts denote long-run and short-run, or permanent and transitory, respectively. Each of these follows its own stochastic process; the long-run component a random walk:

$$y_t^L = y_{t-1}^L + \varepsilon_t \tag{2}$$

and the transitory component a stationary moving average process:

$$y_t^S = \sum_{m=0}^{M} \phi_m \eta_{t-m} \tag{3}$$

where $\phi_0 = 1$. It is assumed that the household is able to separately identify the short and long run components, so that consumption can respond to each. Although the assumption can be criticized, it is not completely implausible. Innovations in y_t^L are immediately consolidated into the income base, while innovations to y_t^S are (at least eventually) transitory, and it is possible that the agent can recognize which is which; salary increases versus consulting income is an obvious possibility, changes in wage scales versus overtime payments is another.

If the permanent income hypothesis is correct, consumption will respond one for one to the change in the long-run component, and by an amount β_t to innovations in the transitory component, where β_t is a function of the real interest rate, the time horizon, and the moving average coefficients in (3),

$$\beta_t = \frac{\sum_{\tau=0}^{T-t} (1+r)^{-\tau} \phi_\tau}{\sum_{\tau=0}^{T-t} (1+r)^{-\tau}}. \tag{4}$$

Hence if f_t is food expenditure, and if the marginal propensity to spend on food is α, then the null hypothesis is that

$$\Delta f_t = \alpha \varepsilon_t + \alpha \beta_t \eta_t. \tag{5}$$

Hall and Mishkin also allow for the presence of transitory consumption, or equivalently for the presence of measurement error in consumption, modelled as a second order stationary MA with coefficients λ_1 and λ_2. Hence, if the transitory component of income is taken to be a third order moving average, we have two equations for the change in income and the change in consumption. For the former,

$$\Delta \tilde{y}_t = \varepsilon_t + \eta_t - (1-\phi_1)\eta_{t-1} - (\phi_1 - \phi_2)\eta_{t-2} - \phi_3 \eta_{t-3} \tag{6}$$

while for the consumption change, and allowing for changes in transitory consumption and consumption measurement error, we have

$$\Delta f_t = \alpha\varepsilon_t + \alpha\beta\eta_t + \nu_t - (1-\lambda_1)\nu_{t-1} - (\lambda_1-\lambda_2)\nu_{t-2} - \lambda_2\nu_{t-3} \tag{7}$$

where ν_t is the innovation in transitory food consumption and β is treated as a constant. The three innovations are assumed to be independent, and to have constant variances, the last being taken as parameters to be estimated along with α, β, the ϕ's and λ's. (These equations are in fact simplified versions of those estimated by Hall and Mishkin, who also allow for the fact that, by the design of the survey, about a quarter of the year has already passed when consumption is measured, so that agents already know something about 'future' income.)

Estimation of these two equations on panel data is straightforward, at least in principle. The data for each of the six years 1969–70 through 1974–75 are used to calculate the cross-sectional sample variances and covariances for the changes in income and the changes in consumption. Theoretical moments are calculated from (6) and (7), and the parameters estimated by a maximum likelihood technique that can be thought of as matching theoretical and sample moments as closely as possible. All the parameters in the two equations are identified. Both moving averages have positive parameters, 0.294 and 0.114 for transitory income, and 0.215 and 0.101 for transitory consumption. The β parameter, which is the response of consumption to innovations in transitory income, is estimated to be 0.292, which can be reconciled with the moving average parameters for transitory income only if the real interest rate is very high, greater than 30%. Alternatively, such an estimate suggests that people have limited horizons, planning over a period that is a good deal shorter than their remaining life-spans, a finding that is consistent with other evidence that consumption track incomes over periods longer than a few years, see particularly Carroll and Summers (1991). The marginal propensity to spend on food is 10%, and the variance of the innovation in transitory income is more than twice as large as the variance in the innovation of the long-run random walk component.

However, there is one major feature of the data that cannot be matched by Hall and Mishkin's model. There is a significant *negative* covariance between changes in consumption and the change in income in the previous period, a covariance which should be zero according to equations (6) and (7). Lagged variables should not be able to predict the change in consumption if the permanent income hypothesis is correct, and the hypothesis founders in much the same way with micro data as it does with macro data, the change in consumption is not orthogonal to the lagged change in income. Hall and Mishkin estimate a second model in which consumers are divided in fixed proportions between permanent income consumers who follow (6) and (7), and liquidity constrained or 'rule-of- thumb' consumers who simply consume their incomes; this model was later adopted by Campbell and Mankiw (1991) to explain their excess sensitivity findings on aggregate data. Hall and Mishkin estimate that the proportion of such 'rule-of-thumb' consumers in the PSID is 20%. Their part of the change in consumption is equal to their change in their income, which is negatively correlated with their lagged income change, so that allowing for such consumers can account for the negative correlation. The findings can also be interpreted in terms of borrowing restrictions, or liquidity constraints. Although it is not true that consumers who cannot borrow will typically consume their incomes, it is true that such consumers who anticipate an income increase will sometimes have to wait to increase their consumption, so that the presence of liquidity constraints can be expected to introduce a correlation between consumption change and predictable changes in income. In the micro data, lagged income change (negatively) predicts income change, and so the presence of consumers who cannot borrow will induce a (negative) correlation between consumption changes and the lagged income changes.

In the macro data, aggregate consumption changes are *positively* correlated with lagged income changes, and lagged income changes *positively* predict income changes, so while the interpretation in terms of liquidity constraints works for both individual and aggregate data, it is hardly possible to claim that these microeconomic results provide any sort of coherent explanation for the macroeconomic findings. Even so, Hall and

Mishkin's results certainly suggest that liquidity constraints might play a useful part in a more complete account of the behavior of consumption.

Reinterpretations: measurement error

Hall and Mishkin allow for the possibility (certainty?) that food consumption is mismeasured in the PSID by including terms for transitory consumption that can also be thought of as reporting error. However, within their econometric framework, and as they note, measurement error in income would destroy the identification of their parameters. The question therefore arises whether all or part of the failure of the permanent income hypothesis in their data can be attributed to this source. The topic has been addressed in a paper by Altonji and Siow (1987). The presence of measurement error in income can hardly be doubted. Altonji and Siow report that the regression coefficient of the change in food expenditure on the change in income is tripled when ordinary least squares is replaced by instrumental variables estimation, using information on lagged income, wages, and other employment information as instruments. The question is what effect the mismeasurement of income can be expected to have on tests of the permanent income hypothesis.

Consider a typical excess sensitivity or liquidity constraints test in which the change in consumption is regressed on the previous periods expectation of the change in income,

$$\Delta c_t = \alpha + \beta E_{t-1} \Delta y_t + u_t. \tag{8}$$

If the permanent income hypothesis is true, then $\beta = 0$, while the presence of some liquidity constrained consumers should show up as $\beta > 0$. Provided there is no measurement error, and provided the change in income is correlated with its lag, then we can use a standard instrumental variables procedure estimating β by

$$\tilde{\beta} = \frac{\text{cov}(\Delta y_{t-1} \Delta c_t)}{\text{cov}(\Delta y_{t-1} \Delta y_t)} \tag{9}$$

and testing whether $\tilde{\beta}$ is zero or positive. However, suppose that the level of income is subject to a white noise error of measurement, so that we observe, not the true change in income, but

$$\Delta y_t^* = \Delta y_t + e_t - e_{t-1}. \tag{10}$$

Substituting into (9), we have

$$\Delta c_t = \alpha + \beta E_{t-1} \Delta y_t^* + \beta e_{t-1} + u_t \tag{11}$$

so that, although the lagged change in true income is orthogonal to the innovation in consumption, the mismeasured lagged change in income, which contains e_{t-1}, is not orthogonal to the compound error in (11) and so is not a valid instrument to consistently estimate β. It appears as if the mismeasurement of income will cause a rejection of the permanent income hypothesis by biasing the estimate of β away from zero. However, this is not the case. If the permanent income hypothesis is true, $\beta = 0$, and there is no bias because the lagged error in (11) no longer appears. Provided the lagged change in income is still correlated with the current change in income, mismeasurement of income cannot explain a non-zero instrumental variable estimate of β if the permanent income hypothesis is true.

It would not be wise to take too much comfort from this result. In particular, the argument ignores measurement error or transitory income in consumption. If, as seems plausible, either measurement error or transitory consumption is positively correlated with the measurement error in income, there will be spurious negative correlation between the consumption change and the lagged income change. Indeed, when Altonji and Siow estimate a (more complicated) version of (8), they find that the permanent income hypothesis is rejected in favor of liquidity constraints when lagged income is used as an instrument, but cannot be rejected if the lagged employment determinants of income are the instruments. Measurement error would seem to be a problem after all.

Reinterpretations: time-series versus cross-sections

Other doubts about Hall and Mishkin's results have been forcefully put in a recent paper by Mariger and Shaw (1990). One issue is whether the original results can be reproduced using other years of data from the PSID. More fundamentally, Mariger and Shaw follow up a remark from Chamberlain's survey on panel data econometrics, Chamberlain (1984, p.1311), that cross-section moments cannot be treated as if they were time-series moments, and argue that it is incorrect to test excess sensitivity by looking at the correlation in the cross-section between changes in consumption and lagged variables. I take up each of the points in turn.

Mariger and Shaw use individual years of the PSID data from 1970 through 1981 (excluding 1972–3) to regress the change in consumption on the change in income lagged once and twice. The coefficients on the second lags are typically small, and typically smaller than their standard errors, and can be ignored for the moment. The coefficients on the first lag, which correspond to Hall and Mishkin's excess sensitivity test, are negative in 1971, 1974, 1975, 1976, and 1980, the first three of which are used by Hall and Mishkin, but is positive in 1970, 1977, 1978, 1979, and 1981, only the first of which appears in the Hall and Mishkin sample. None of the positive estimates are larger than their standard errors, and all of the negative ones are, except for 1980. On balance, one might conclude with Hall and Mishkin, that there is more evidence of a negative covariance than of a zero one. Indeed, the pooled estimate for the whole period shows a negative coefficient of –0.0057, but the standard error is 0.0053, compared with an estimate of –0.0181 (0.0101) over the three years 1971, 1974, and 1975, all of which are in Hall and Mishkin's sample. The result that seemed to indicate liquidity constraints does not seem to be a feature of every year's data from the PSID, and it does not seem to be present at all in the more recent data.

Mariger and Shaw suggest an interpretation of this evidence that explains the year to year variability in the PSID covariances, as well as why estimation over a short period might lead to incorrect inferences. The point can be made most clearly using a very simple model of individual income processes. Suppose that aggregate income is a random walk, that everyone gets a share of this aggregate, together with an idiosyncratic and transitory income shock. Household i therefore experiences an income change that can be written as

$$\Delta y_{it} = \varepsilon_t + u_{it} - u_{it-1} \tag{12}$$

where ε_t is the innovation to the random walk and is common to all, and u_{it} is idiosyncratic transitory income. I assume that these individual components of income are uncorrelated across individuals, and that the aggregate shock ε_t is known to everyone, an issue to which I return in Section 2 below. Suppose too for simplicity that horizons are infinite so that the change in consumption for each household is given by

$$\Delta c_{it} = \varepsilon_t + \frac{r}{1+r} u_{it} \tag{13}$$

We are now in a position to calculate the regression coefficient of the change in consumption on the lagged change in income. Suppose we do this for a single year t. Elementary calculation gives:

$$\operatorname*{plim}_{n\to\infty} \tilde{\beta}_t = \operatorname*{plim}_{n\to\infty} \frac{\sum_{i=1}^{n} \Delta c_{it} \Delta y_{it-1}}{\sum_{i=1}^{n} \Delta y_{it-1}^2} = \frac{\varepsilon_t \varepsilon_{t-1}}{\varepsilon_t^2 + 2\sigma_u^2} \tag{14}$$

where the sums are taken over the n households in the cross-section. Although the time average of the β_t estimates will converge to zero as the number of time periods goes to infinity, that is not the situation that we are in. For single years, or for short panels, the coefficients will vary randomly from year to year, depending on the aggregate shock for each year. Indeed Mariger and Shaw, using a random coefficients model, argue that the pattern of coefficients in the data are consistent with this interpretation, and that there is therefore no evidence for liquidity constraints. They also attribute Altonji and Siow's positive findings to the fact that they use a longer data period than do Hall and Mishkin, so that there is less scope for the average coefficient to deviate from zero.

In the particular example given above, where the aggregate shocks are simply added to the idiosyncratic components of income, the result depends on computing the regression without an intercept. The means of the consumption change and lagged income change are ε_t and ε_{t-1} respectively, so that regressions that allow dummies for each period would not be expected to show excess sensitivity, even with few time periods. It should also be noted that Altonji and Siow, unlike Hall and Mishkin, used time dummies in their preliminary regressions, so that their income and consumption changes have cross-household means of zero by construction. However, there are plausible specifications for individual incomes where time dummies will not remove the effects of the aggregate shocks. For example, add an interactive term to (12) to give

$$\Delta y_{ti} = \varepsilon_t + \eta_i \varepsilon_t + u_{it} - u_{it-1} \qquad (15)$$

so that the effect of the aggregate shock varies over individuals according to a time-invariant parameter η_i which I assume to have mean zero over the population. According to (15), although everyone's income has a component that goes up or down with the aggregate economy, each person's total income is affected differently, some benefiting more than average when the economy is booming, some less, and some losing out. The change in consumption will be as in (13) except that the common shock ε_t must now be multiplied by the idiosyncratic factor $1+\eta_i$.

We then follow Altonji and Siow and remove time effects in preliminary regressions before running the excess sensitivity tests in the cross-section. This amounts to subtracting from the consumption and income changes the cross-sectional means for each year $\Delta c_{.t}$ and $\Delta y_{.t}$. The probability limit in (14) is now replaced by:

$$\operatorname*{plim}_{n\to\infty} \frac{\sum_{i=1}^{n} (\Delta c_{it}-\Delta c_{.t})(\Delta y_{it-1}-\Delta y_{.t-1})}{\sum_{i=1}^{n} (\Delta y_{it-1}-\Delta y_{.t-1})^2} = \frac{\varepsilon_t \varepsilon_{t-1} \sigma_\eta^2}{\varepsilon_t^2 \sigma_\eta^2 + 2\sigma_u^2} \qquad (16)$$

so that the removal of the time means has not solved the problem. Of course, if we knew in advance how the aggregate shock affected each individual, we could design an estimator that was consistent in the cross-section, but it is not obvious how such knowledge could be obtained, so that it is hard to see how we can believe that the aggregate shocks have been eliminated. If so, it would appear that excess sensitivity tests on panel data require at least enough time periods to allow some assessment of the variability over time of the excess sensitivity parameter.

Other evidence from the PSID

Two other recent studies of food consumption base their analysis on the Euler equation for intertemporal optimality rather than the permanent income (quadratic felicity) approach of Hall and Mishkin. Zeldes (1989) and Runkle (1991) both employ use isoelastic (constant relative risk aversion) felicity functions which yield, as an approximation, that the rate of growth of consumption is a linear function of the expected real interest rate. If household specific variables are also included, we can write for household i at time t:

$$\Delta \ln c_{it+1} = \alpha + \eta_i + \zeta_t + \beta_1 E_t r_{it+1} + \beta_2 age_{it+1} + \beta_3 \Delta \ln FS_{it+1} + z_{it} + \varepsilon_{it+1} \tag{17}$$

where η_i and ζ_t are individual and year effects respectively, *age* is the age of the household head, *FS* is a measure of family size, and z_{it} is a positive number if liquidity constraints are binding, and is zero otherwise. Equation (17) is Zeldes' specification, but it is also useful for describing Runkle's results.

Family size, age, and age squared can be expected to appear in the felicity functions, and so their first differences are included contemporaneously with the change in consumption. The family fixed effects capture other household specific determinants of consumption change, and the time effects any additive aggregate shocks in individual household innovations. The liquidity constraints variable is derived by Zeldes from the Euler equation when borrowing is not allowed and is interpreted as follows. For a household that is not constrained *in the current period*, the Euler equation will hold, even if liquidity constraints can be expected to bind at some point in the future. For a household that wishes to borrow but cannot, consumption today is too low relative to tomorrow's consumption, so that one period ahead consumption growth will be too high, which is represented in the equation by a positive z_{it}. Liquidity constraints cannot generate too low a growth rate, because the household, even if it cannot consume more, can always save more, and thus increase its growth rate of consumption. Zeldes' methodology explicitly recognizes this behavioral asymmetry induced by borrowing constraints, and represents a marked improvement over assuming that liquidity constraints cause consumers simply to spend their incomes. On the other hand, as demonstrated in Deaton (1991a), liquidity constraints can have a dramatic effect on consumption with only occasional violations of the Euler equation, so that tests based on the Euler equations may not be very powerful.

Zeldes uses data from the PSID for years from 1968 through 1982, so that, deleting years without food expenditure, he has up to 10 observations for each household; in practice the average is between three and four, depending on the experiment. Each household is allocated in each year to a high asset or low asset regime, depending on its ratio of assets to income, and variants of (17) are estimated for each subsample separately. The interest rate variable is instrumented by each household's marginal tax rate in period t; without variations in tax rates the interest rate effects would not be distinguishable from the year effects ζ_t. In the regressions, the year effects dealt with using dummy variables, and the household fixed effects are eliminated by sweeping out the individual household means. This differencing introduces a time average into the error term, which therefore can no longer be guaranteed to be orthogonal to variables dated t and earlier. Nevertheless, if fixed effects are not taken into account, but are in fact present, they will generally be correlated with assets, so that splitting the sample by asset levels will cause selection bias. For example, especially cautious households will have both high assets and high rates of consumption growth. If (17) is estimated on the high asset sample, where z_{it} should be zero, and the parameters used to calculate predicted growth rates of consumption for the low asset sample, the average underestimation is 1.7%

a year, with a t-value of 1.63. Supporting evidence comes from estimating (17) for each sample without the liquidity constraint variable, but including $\ln y_{it}$. This excess sensitivity test generates a significant negative coefficient for the low asset group; the lower is income, the greater the likelihood of not being able to borrow, and the higher the growth rate of consumption. For the high asset group, the coefficient is also negative, but half the size and statistically insignificant. Note that, although both these coefficients may be biased downwards by the treatment of the fixed effects, there is no reason to suppose that the bias is any worse for the low than for the high asset group, and there is no reason to challenge Zeldes' conclusion.

Although Runkle also uses the PSID, with data from 1973 to 1982, his results are quite different. His version of (17) includes only the interest rate and age terms, although he also investigates whether time and individual effects should be included. He detects neither of these in the data, although there is evidence of substantial measurement error in consumption, which is allowed for by adding a moving average measurement error term to (17), and making appropriate corrections to standard errors and test statistics. Again the technique is instrumental variables, but now using generalized methods of moments estimation, with the overidentification test statistic as a measure of model adequacy, and of orthogonality between the instruments and the innovation. If the test fails, at least some of the instruments, all of which are dated t or earlier, cannot be orthogonal to the supposed innovation, so that we have the equivalent of excess sensitivity. The instruments are a constant, age, the marginal tax rate, hours worked, disposable income, asset income, and liquid assets in both $t-1$ and $t-2$. Runkle accepts the hypothesis for the sample as a whole, as well as for subsamples according to asset to income ratios, evidence that he takes to be strongly supportive of the theory. Unlike Zeldes, he also finds strongly significant interest rate effects, although the estimated coefficient is around 0.45 in both studies.

I find it difficult to reconcile Runkle's results with Zeldes' findings, or indeed with those of the other studies of the PSID that I have already discussed. For example, in one experiment Runkle includes time dummies as instruments, but still obtains an overidentification test statistic that is insignificant, and interprets the result as demonstrating that aggregate shocks are not important. If so, the rate of growth of consumption in the aggregate data should be constant, something that we know is not true. Similarly, it is surprising that none of hours, income, or liquid assets affect the test statistic, given the results of Hall and Mishkin, Zeldes, and even Mariger and Shaw on the lagged income change, and those of Hotz, Kydland and Sedlacek (1988) that lagged hours strongly influence food consumption. The change in consumption in Runkle's subsample seems to be orthogonal to everything except age and the marginal tax rate!

I suspect that, as noted by Runkle, a major source of discrepancies is the criteria used to select the sample. Zeldes has almost *six times* as many data points as does Runkle, who has an average of little more than one change per household, 2,830 observations on 1,144 households. Given the difficulties of working with data sets such as the PSID, which after all was designed for quite different purposes, and given the immense amount of work that would be involved, it is not surprising that authors do not attempt to replicate previous work when proposing their own models. However, it is clear from other work using the PSID, particularly studies of labor supply, that results can be sensitive to the construction of the subsample, and the issue needs more investigation. Ideally, we need a study corresponding to that of Mroz (1985) on labor supply, where the specific contributions of sample selection, econometric technique, and model choice are carefully disentangled.

Other tests: the US, Italy, Japan, Norway, and Côte d'Ivoire

The PSID is not the only source of data that can be used to test the theory of consumption. Hayashi (1985a, 1985b) has pioneered the use of imaginative tests on quite different data. In Hayashi (1985a), use is made of a single cross-sectional household survey, The Survey of Financial Characteristics of Consumers, collected in the United States in 1963–4 by the Board of Governors of the Federal Reserve System. The basic idea is again to split the sample into those who are possibly liquidity constrained and those who are almost certainly not, to estimate a consumption function for the latter, and see how it predicts the former. Hayashi selects the unconstrained group according to its saving rate; households who are saving positive amounts are not up against a borrowing constraint. Consumption is regressed on assets, income, age, and interaction terms, and a Tobit procedure used to correct for the bias that would otherwise result from selecting the sample on the endogenous variable. The estimated parameters can then be compared, using a Hausman test, with the estimates obtained from the whole sample, which if the theory is correct, should have the same probability limit, but be more efficient. As it turns out, the parameters are not the same, and the consumption predictions for the low asset group using the parameters obtained from the high asset sample tend to overpredict consumption of the former, especially for younger households, who are those whom we might expect to be most likely to be subject to borrowing constraints. These (earlier) results are, of course, very similar to those obtained by Zeldes. They have also been replicated in a matching study by Jappelli and Pagano, who use an Italian cross-section from 1984. Once again the results indicate that the largest shortfall in desired consumption is for households headed by people under 30. The magnitudes of the shortfalls are proportionally larger than in Hayashi's estimates, which Jappelli and Pagano interpret as reflecting the relatively more developed system of consumer credit in the United States compared with Italy, even given the fact that the Italian survey is 20 years later than that from the United States.

Hayashi (1985b) has also looked at consumption behavior in a short panel of Japanese households. In this remarkable survey, collected in 1981–82, households were visited four times each at quarterly intervals, and asked, not only about their consumption and incomes, but also about their expectations of consumption and income in the following quarter. The availability of this direct information on expectations means that Hayashi is uniquely able to avoid the time-series cross-section problem, and use estimation techniques that allow for the presence of quite arbitrary aggregate shocks. He works with a sevenfold disaggregation of consumers' expenditure, and permits each to have some degree of durability. The model is the one originally proposed by Mankiw (1982), whereby felicity depends, not on purchases, but on the accumulated stock of the good. Marginal utility is therefore also a function of the stock, so that if we adopt the permanent income assumptions of certainty equivalence, and a constant real rate equal to the rate of time preference, it is the stock, not the purchases, that follows a random walk. Hence if the change in stock is an innovation ε_{t+1}, and purchases in $t+1$ are the difference between the stock at $t+1$ and $(1-\theta)$ times the stock at t, we have at once that,

$$\Delta c_{t+1} = u_{t+1} - (1-\theta)u_t. \tag{18}$$

Hayashi adds additional stochastic terms to (18), to allow for preference shocks in the two periods, and to incorporate measurement error in the reported expenditures. He also follows Hall and Mishkin by adding a fixed proportion of rule-of-thumb consumers who consume their incomes. These equations can then be used to derive theoretical covariances between the expected and unexpected changes in consumption and the expected and unexpected changes in disposable income. If there is no measurement error in income, knowledge of these covariances is sufficient to identify the parameters of the model without having to assume that the individual innovations have zero mean, or are uncorrelated in the cross section with lagged variables like income. Instead these covariances are estimated in each period along with the other parameters of the model. In order to minimize mismeasurement of income, Hayashi excludes all but wage earners from his sample.

Except for food, Hayashi finds some durability for all the commodity groups. He also finds evidence of excess sensitivity, with the proportion of rule-of-thumb consumers estimated to be 15%. Hayashi also runs standard excess sensitivity tests for food, where there is no complication from durability, by regressing for each quarter in the cross-section the change in food against the anticipated and unanticipated changes in disposable income. These coefficients are 0.014 (0.004), 0.015 (0.005), and 0.025 (0.006) for the three quarterly differences, all small, but all significantly positive. Of course, these results, unlike the earlier ones, could be explained by an appropriate pattern of aggregate shocks, but they are also in accord with the evidence for liquidity constraints from the PSID, and they come from data where the presence of excess sensitivity has been established by tests that are robust against the aggregate shock problem.

Durability by itself can account for a negative correlation between the change in consumption change and the lagged income change. This result can be seen from equation (18) where Δc_{t+1} is affected by the lagged innovation, which is itself typically correlated with lagged income. When felicity functions depend on stocks, so that the stocks follow a random walk, an innovation in income in t will increase stocks in t, implemented by making purchases in t, but since the stock in t+1 is expected to be the same, expected consumption next period will be confined to making good the physical deterioration. Consumption can therefore be expected to fall in the period after a positive income innovation. Of course, this is not a issue for food, where stocks and purchases will usually be very similar. However, it is notable that Hayashi finds evidence of excess sensitivity even after allowing for the effects of durability, an allowance that would typically reduce the role of excess sensitivity in explaining the negative correlation in the micro data.

Flavin (1991) uses a previously unexploited data set, a subset of 1600 households from the 1967 Survey of Consumer Finances who were reinterviewed in the two succeeding years. These data do not measure consumption directly, but there is information on income and on a detailed menu of assets. Flavin regresses the change in savings on the change in income, using an instrument constructed from the reported future income expectations of the households in the panel. She finds that only 20–30% of anticipated income changes are saved, as opposed to the 100% predicted by the permanent income hypothesis. However, when households are split by wealth status, there is no suggestion that the excess sensitivity is any less severe for rich households than poor ones, so that it is far from clear that the findings can be attributed to the operation of liquidity constraints.

Mork and Smith (1989) have tested a model similar to Hall and Mishkin's using panel data from Norway. The data come from two separate two-year panels in 1975–76 and 1976–77, and respondents also reported consumption and income in the year prior to the surveys, so that there are three successive observations for each of the households. However, as with the PSID, there are problems matching the timing of reported consumption and reported income, so that changes in consumption will generally be correlated with the lagged income changes, even if the permanent income hypothesis is correct. Mork and Smith therefore base their orthogonality test on the independence of the change in consumption and the *level* of income two periods previously. Since income in t–2 is certainly known at t, it can be argued that it should be orthogonal to consumption change in period t. In the data, the correlation is negative, but it is not significantly different from zero, and Mork and Smith conclude in favor of the hypothesis. However, it is not clear that such a test is likely to be very powerful in detecting the effects of liquidity constraints, if indeed they are present. Liquidity constraints will induce a relationship between the change in consumption and the anticipated change in income, but there seems little to reason to suppose that the latter will be strongly related to the previous *level* of income.

In my own work, Deaton (1991b), I have looked at a very different set of data, from households in Côte d'Ivoire. The data come in two separate panels, 1985–86, and 1987–88, but there are only two annual observations for each household. In such circumstances, it is difficult to say very much about consumption unless strong assumptions are made. However, Côte d'Ivoire is a country where there has been very little real economic growth over the last twenty years, so it makes some sense to analyze behavior under the assumption that individual income processes are stationary, so that trend issues do not arise. One can then visualize these predominantly agricultural households, many of whom are tree-crop farmers, trying to smooth their consumption in the face of the quite large income fluctuations generated by year to year variations in weather, pests, and fires.

The basic econometric problem in working with these data is how to estimate any sort of dynamic model, while allowing for at least minimal individual heterogeneity in income processes. The trick is to use the implications for saving of the assumption that incomes are stationary. In particular, consider Campbell's (1987) saving equation, whereby under the permanent income hypothesis, saving is the discounted present value of expected future declines in income. Hence, if income is stationary, the *unconditional* mean of saving is zero for each household, no matter what its average income level. I have assumed away any life-cycle motives, and saving by these households is only to smooth out their income fluctuations; there is no reason for them to accumulate assets over the long term. Suppose too that each household's income y_{it} is stationary around its own individual mean f_i. Then consumption must also be stationary around the same individual mean, because the mean of saving is zero for everyone. An appropriate excess sensitivity test, accounting for fixed effects, would be to run the regression

$$c_{it} - f_i = \gamma + \alpha(c_{it-1} - f_i) + \beta(y_{it-1} - f_i) + \varepsilon_{it} \tag{19}$$

and test that $\alpha = 1$ and $\beta = 0$. Unfortunately, there is no way to eliminate the unobservable fixed effects from this regression without at least three observations, so that, at the least, I could difference one more time. *Faute de mieux*, I can simply run the regression of consumption against lagged consumption and lagged income, absorbing the fixed effects into the error term, so that

$$c_{it} = \gamma + \alpha c_{it-1} + \beta y_{it-1} + \varepsilon_{it} + \big((1-\alpha) + \beta\big) f_i \tag{20}$$

a regression which, in general, will deliver inconsistent estimates of both parameters. However, if the permanent income hypothesis is true, the coefficient multiplying the fixed effect in the regression is zero, so that there is no bias or inconsistency. If α is not unity, and β is not zero, the permanent income hypothesis cannot be true, fixed effects or no fixed effects. The argument is the same as that for testing the hypothesis with mismeasurement of income, see the discussion surrounding (11) above.

As with all the other studies, it is necessary to make some attempt to allow for measurement error. I do this by using the various correlates of income as instruments; areas of various crops, hours worked, and weather conditions, none of which are used directly in the calculation of income and consumption. It is easily checked that moving from OLS to IVE does not affect the unimportance of the fixed effects if the permanent income hypothesis is true. However, my estimates do not support the model, at least in this form. Although the data are consistent with the hypothesis that β is zero, they are not consistent with α being unity, although in the presence of the fixed effects, and although I know that α is not unity, I cannot obtain a consistent estimate of it. Habits or slow adjustment are an obvious possibility, but so must be the suspicion that the effects of measurement error have not been completely eliminated by the instrumentation. However, there is no evidence here for the standard excess sensitivity story, although in view of the extreme difficulty of estimating income for self-employed farmers, the results should be treated with a great deal of caution.

A tentative summary

It is difficult to distill any very firm conclusions from all of this evidence, which I have presented in some detail so that it is possible to see the diversity of the results, as well as the difficulties that have to be overcome in using the microeconomic data on consumption. Perhaps it is not reasonable to expect uniform results from different data sets, but as we have seen for the PSID, widely divergent results can be obtained even from the same data. Nevertheless, it is would be safe to conclude that the evidence against the theory in the micro data, if it exists at all, is weaker and less transparent than in the aggregate data. There is less evidence against the permanent income hypothesis for individual agents than there is against the permanent income hypothesis for the representative agent. Even so, my own judgement is that the micro data provide *some* evidence against the hypothesis, and that the problems are those that would be expected if borrowing constraints exist for at least some of the consumers some of the time. While I have no doubt that all of the contrary evidence can be explained if we try hard enough, I find the evidence for some form of liquidity constraints both plausible and convincing. Hayashi's two studies offer good evidence for such effects for both the United States and Japan, as does Jappelli and Pagano's replication for Italy. As far as the PSID is concerned, although it must be recognized that the negative correlation between the change in food consumption and the previous change in income is neither constant over time, nor free of contamination by possible aggregate shocks, it is present in most years, and it is in the direction predicted by the presence of liquidity constraints. Furthermore, it is consistent with Zeldes' evidence that low asset consumers are those mostly affected, evidence that I find generally credible and that is robust against at least the simplest type of aggregate shocks.

It should also be emphasized that the standard tests on panel data, whether for violation of the Euler equations, or for excess sensitivity, may not be very good at detecting the operation of liquidity constraints. In Deaton (1991a) I construct an example of an impatient consumer, who would ideally like to borrow for high consumption early in life, but whose consumption plan is forced to be a stationary process by a prohibition on borrowing. Nevertheless, the Euler condition for optimal intertemporal allocation is satisfied in most periods of life. As we shall see, liquidity constraints have the effect of reducing the effective planning horizon, so that it is also worth recalling Hall and Mishkin's finding that the horizon is only a few years ahead even for their permanent income consumers, as well as Hayashi's (1982) result that much of the time series evidence can be reconciled with the theory if labor income is discounted at a much higher than market rate.

To the results of the econometric tests from the panel data must be added other, less formal evidence. In particular, the findings that consumption tracks income closely over the life-cycle, over occupations, and over countries, should also be taken into account. Since we know that the occupational and national profiles of consumption and income are reasonably stable over time, it is possible to use future income profiles to predict that the consumption of someone aged 30 will grow by less over the next two decades if the person is a car mechanic rather than a trainee physician or lawyer, or if the person is an Ivorian as opposed to a Korean or a Thai. Consumption change is related to predictable income change, as it should not be if the theory is correct. That such tracking does not appear in the tests on the PSID or other data sets may be more a reflection of the short term nature of these tests rather than the absence of the phenomenon.

There is also some direct time-series evidence that is not subject to the ambiguities of the excess sensitivity tests in Chapter 3. From the end of post-war rationing until the late 1960's, the British government controlled the terms on which consumers could borrow for the acquisition of durable goods. The controls were specific by type of durable good, with distinctions between cars, radio and electrical goods, and furniture and floor coverings. The regulations, which applied only to 'hire-purchase finance companies' and not to banks or other financial institutions, defined the minimum downpayment required as well as the maximum number of months over which the loan had to be repaid. As with all such controls, the market eventually found ways to undermine their effectiveness, but for many years, changes in the percentage downpayment exerted a strong influence on total consumers' expenditure, and was used by the Treasury as a fine-tuning instrument of macroeconomic control. See Dow (1964, pp. 246–8, 278–82) for an institutional description, and Stone (1966, 1973) for the (very convincing) evidence of the effects on consumption. Credit restrictions, even over a small segment of the market, can clearly have a large effect on consumer behavior. In the United States, Wilcox (1989) has found overwhelming evidence that payments of social security benefits, even when announced many weeks in advance, are accompanied by increases in expenditure when the checks arrive, and not when the increase in benefits is announced. Of course, in this case, people may not believe what they read, or they may not choose to keep themselves informed, an issue to which I shall return below. The 'knowledge' interpretation is given some further support by Wilcox's (1987) other finding, that income tax refunds are typically *not* associated with synchronous increases in consumption.

Finally, I see no reason to neglect the informal evidence that we see in the environment around us, that tells us (or at least me) that I do not have access to unlimited borrowing, and that people who are poorer than I am have even less. A young, temporarily poor, but impatient consumer, who expects income to grow in the future, would want to have a large negative net worth early in life. I see no reason to not to believe that there are many such people, and I do not believe that they can borrow as much as they would wish. Indeed, Jappelli (1990) reports that in the US Survey of Consumer Finances, 12.5% of households reported in 1982 that a financial institution had turned down their request for credit in the last few years, with a further 6.5% reporting that they had been discouraged from applying by the belief that their request would be denied.

It is important not to misinterpret the seriousness of this evidence against the theory. If it is true that some people would like to borrow and cannot, that does not mean that there are many other households, perhaps most, who either can borrow, or who are capable of detaching their consumption streams from their incomes without no or only limited access to credit. That the permanent income story is not *all* that is going on does not mean that it is not a great deal of it. There is nothing in the evidence of this section that suggests that everyone spends his or her income, that the dynamics of income and consumption are identical, or that denies the basic insights of the theory of intertemporal allocation.

2 The reconciliation of micro and macro evidence

If we take the view that the household data are basically supportive of the theory, then it is a puzzle why the theory is rejected in the aggregate. Alternatively, if we accept the arguments in favor of the (at least occasional) importance of liquidity constraints, there is no surprise that the theory should be rejected in aggregate, but there is still a problem in reconciling the precise nature of the macroeconomic and microeconomic evidence. The time series properties of aggregate income and consumption appear to be quite different from their macroeconomic counterparts, and the predictability of consumption changes in the two environments seems to coincide mainly by chance. The household data usually show negative autocorrelation for both income and consumption changes, so that a negative correlation between consumption and lagged income changes can be interpreted as predictable income changes being positively correlated with changes in consumption. In the aggregate data, both consumption changes and income changes are *positively* autocorrelated, so that the *positive* correlation between consumption change and lagged income change can be interpreted, as before, as showing that consumption change can be predicted by predictable income growth. Given the differences in all but the final conclusion, such results can hardly be regarded as a triumph for a representative agent interpretation of the data.

If the consumption of each individual agent follows a martingale, then there are well defined conditions under which mean consumption will do likewise. If $c_{it+1} = c_{it} + u_{it+1}$, then we have that $c_{t+1} = c_t + u_{t+1}$, for the corresponding averages. However, in taking the mean, it is implicitly being assumed that the same individuals are present in the two periods, so that there can be neither births nor deaths. Secondly, while the individual u_{it+1} terms are innovations, they are unpredictable only with respect to the information sets of each individual, so that there is no presumption that the average u_{it+1} will be an innovation with respect to anything. However, as pointed out by Grossman and Shiller (1982), if macro variables are known to everyone, each individual innovation will be orthogonal to the macro information, and so will be their mean. Hence, if people live for ever, and if they are well-informed about macroeconomic variables, aggregate consumption will follow a martingale if individual consumptions do so. It is the possible failure of these two conditions that we have to explore as potential sources of the differences between the micro and macro findings, and I consider each in turn.

In this section, I work under the basic assumption that the theory is true for individual data and explore the consequences of failure of the aggregation assumptions. The analysis of aggregation when people are liquidity constrained is a good deal more difficult, although a first numerical attempt is reported in Deaton (1991a). Note also that what I am concerned with here is the aggregation of the martingale property, that there exists some function of consumption whose future change is unpredictable given current information. I shall not be concerned with the prediction that relates consumption growth to the rates of return and risk properties of financial assets since I do not believe that this supposition is well-based even at the microeconomic level. Perhaps a fifth of consumption in the United States is attributable to consumers who do not even have a checking account, let alone interest bearing assets, so that it stretches credulity to model their consumption decisions on the assumption that they can borrow on the same terms as the US federal government. Even if this chimera were mistaken for reality, there is no reason to suppose that it would aggregate to the macro level unless, against all available evidence, people are immortal.

Aggregation with finite lives

A world populated by finitely-lived life-cycle consumers will generate aggregate consumption that exhibits both excess sensitivity to current income and excess smoothness to innovations in current income. This result is due to Clarida (1991) and to Galí (1991), and my discussion is based on their work. I take the simplest possible case of Clarida's model. Each worker lives for three periods, working in the first two periods of life, and retiring in the third. In year t, each worker receives an identical amount of labor income y_t while they are working, and zero in retirement. The common quantity y_t follows a random walk with drift g, so that average labor income, which is two thirds of what each worker receives, also follows a random walk with drift two thirds of g. Suppose that the interest rate is zero, and that everyone is a pure permanent income life-cycler. Now compare the situation in t with that in $t-1$ and look at the consumption of each group. Those who are now retired but were old workers in the last period receive no income, which is what they expected, so their consumption does not change. The old workers now, who were previously young workers, have had an innovation ε_t, half of which they will consume now, so that their consumption change is $\varepsilon_t/2$. The new born workers receive y_t, they expect y_t+g next period, and so will consume $(2y_t+g)/3$ in period t. The picture is completed by looking at the consumption of the newly dead group, previously retired. In the first period of their lives, which was $t-3$, they consumed $(2y_{t-3}+g)/3$, in middle age, period $t-2$, this plus $\varepsilon_{t-2}/2$, and the same in the last period of their lives, $t-1$, since there are no surprises after retirement. But

$$y_{t-3} = y_t - 3g - \varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-2} \tag{21}$$

so that, if all the changes in consumption are added up, we get

$$\Delta C_t = 2g + \frac{7}{6}\varepsilon_t + \frac{2}{3}\varepsilon_{t-1} + \frac{1}{6}\varepsilon_{t-2}. \tag{22}$$

Recall that average income is two-thirds of total income, with corresponding drift and innovation, so that if (22) is rewritten in terms of the change in average consumption, and the drift $\tilde{g}$ and innovation $\tilde{\varepsilon}_t$ in average income, we obtain finally

$$\Delta c_t = \tilde{g} + \frac{7}{12}\tilde{\varepsilon}_t + \frac{1}{3}\tilde{\varepsilon}_{t-1} + \frac{1}{12}\tilde{\varepsilon}_{t-2}. \tag{23}$$

The change in average consumption has a drift term equal to the drift in average income, and is a distributed lag of the innovations in average income with weights adding to one. Consumption is not orthogonal to lagged innovations, nor does it respond one for one to innovations in current income, even though the latter is a random walk. Aggregation generates both excess smoothness and excess sensitivity although there is none for any individual consumer. Clarida's paper works out the more general case with arbitrary lengths of working and retirement periods, and with a positive real rate of interest, and shows that the result still holds, although the numerical values of the coefficients will be different. There are two factors driving the result. First, since people receive income only when they are working, and have to make provision for retirement, they will consume only a fraction of an innovation to income, even when they know that the innovation will persist through the rest of their working lives. At the aggregate level, the response of consumption to aggregate innovations is essentially the average of these fractional responses, and will be a good deal less than unity. Second, because people die and are replaced by new borns, who do not share their history, the income experience of the now dead generation affects the change in consumption. In addition, because incomes are growing, the prospects for the new generations are better than their defunct forebears, so that there is an upward drift in consumption. It might be thought that these generational effects would be small over short time periods such as a quarter or a year, but working through the arithmetic shows that this is not so. Indeed, the model can be reworked in continuous time with very similar results.

Another feature of these results has been emphasized in Galí's work. The aggregate equation (23) implies that average consumption and aggregate *labor* income are cointegrated, and are tied together in the long run. This can be seen in the example either by subtracting the average income change from the consumption change in (23) and noting that the resulting moving average on the right hand side involves only the first differences of the innovations, or by aggregating the model explicitly to recover average consumption, and showing that it differs from income by a quantity that is stationary.

The permanent income hypothesis implies that, for individual agents, consumption and *total* income, including property income, are cointegrated, see for example Campbell (1987), but that there is no long-run relationship between consumption and labor income at the microeconomic level. Indeed, the absence of such a relationship is at the core of the theory of intertemporal allocation without borrowing restrictions. In the aggregate, the turnover of generations transmits income growth to consumption growth. The difference between consumption and income is stationary and we avoid the absurdity of having to believe that in aggregate, desired consumption could be many times larger than actual income, as well as the implication of equation (26), that saving will be large and negative when labor incomes are expected to grow.

There remains the empirical question of whether an aggregate model, modified to take these considerations into account, can account for the failures of the theory in the aggregate data. Both Clarida and Galí conclude not. Clearly, the direction of the effects is right, and they go someway to explaining both excess smoothness and excess sensitivity. Clarida calculates that for reasonable values of the interest rate and of the length of life and retirement, but still on the assumption that income processes are the same, the warranted change in aggregate consumption to an innovation in random walk income should be around two-thirds, still too large to account for the actual smoothness of the aggregate series. Similarly, it is clear from Galí's work that the cointegration between aggregate consumption and aggregate labor income, although real enough, is a long-run phenomenon, and cannot account for the short run tracking of consumption and income that generates the excess sensitivity results.

Individual households and aggregate information

The leading candidate left to account for aggregation problems is the failure of individuals to be aware of aggregate information. The issue here is partly accessibility, since consumers may find out about aggregate income or other macro variables only with a substantial lag, and partly relevance, whether aggregate information is sufficiently valuable to individual agents so that they will bother to obtain it, even when it is available. Large corporations buy economic forecasts, but private individuals rarely do so. Even if the information is readily available in newspapers or government publications, many people would require education to interpret it and to infer from it the implications for their own futures. Indeed, when many graduate students in economics do not know the gross national product of the United States within an order of magnitude, it is strange to suppose that those with no training in economics can monitor and interpret the latest macroeconomic shock. The consequences of aggregate information with lags, and of no aggregate information, have been worked out in papers by Goodfriend (1988) and Pischke (1991). Here I follow the simplest example in Pischke, which neatly highlights the issues.

Suppose that average income follows a random walk with drift, and that each consumer's income is the average plus a idiosyncratic component that is purely transitory, represented by white noise. The first difference of individual income is therefore the first difference of the random walk, including drift, plus the first difference of white noise. Hence:

$$\Delta y_{it} = g + \varepsilon_t + u_{it} - u_{it-1} \tag{24}$$

where the suffices differentiate the aggregate, common components from those that are specific to the individual households. Pischke's assumption is that the household does not choose to acquire the aggregate information about the macro shock that would allow the separation of the two components in (24). Instead, each person observes only their sum, which is the moving average process:

$$\Delta y_{it} = g + \eta_{it} - \lambda\eta_{it-1} \tag{25}$$

The parameter λ can be calculated by solving the quadratic that results from equating the autocorrelation coefficients of the original and derived processes, (24) and (25). However, we know from studies of the PSID and elsewhere that aggregate shocks account for very little of the variation in individual incomes, so that the parameter λ must be close to unity. Individual incomes are much closer to white noise than to random walks.

It is now straightforward to calculate individual and aggregate consumption change. Again, suppose that each household satisfies the infinite life permanent income model so that we abstract from the finite life aggregation phenomena of the previous subsection. For a pure permanent income consumer, the warranted change in consumption is

$$\Delta c_{it} = (1 - \frac{\lambda}{1+r})\eta_{it} \tag{26}$$

so that, since λ is close to unity, individual consumption is much smoother than the individual income changes. Although the consumer knows that there is an aggregate persistent shock in his or her own innovation, there is no way of knowing exactly what it is, so the response of consumption is only a little more than it would be if the income shock were purely transitory. The change in aggregate consumption is obtained by averaging (26) over the population, so that, dropping the individual suffices so as to denote means,

$$\Delta c_t = (1 + \frac{\lambda}{1+r})\eta_t \tag{27}$$

However, if we compare the population means of (24) and (25), and use the fact that the idiosyncratic components have zero mean over the population, we have

$$\varepsilon_t = \eta_t - \lambda\eta_{t-1}. \tag{28}$$

Hence, from (26), the change in aggregate consumption follows the autoregressive process

$$\Delta c_t = \lambda\Delta c_{t-1} + (1 - \frac{\lambda}{1+r})\varepsilon_t. \tag{29}$$

Finite life effects would add an intercept term to (29), as well as some additional terms in lagged innovations, although they will not be as important as in Clarida's example where all incomes were identical.

Far from being a random walk, the change in aggregate consumption is strongly autoregressive, with only a small response to the innovation in aggregate income. If (29) really holds in the data, aggregate consumption will indeed be very smooth, responding only with long lags to innovations in aggregate income. Although this is no more than an illustrative example, it captures a number of features of reality. Data from the PSID, MaCurdy (1982) and Abowd and Card (1986), as well as from the Survey of Income and Program Participation, Pischke (1991), suggest that the income changes of individual households follow an MA(2) rather than an MA(1), although the latter is a relatively good approximation. Presumably, the idiosyncratic elements of individual incomes also have some persistence over time, which would tend to increase the λ parameter in the example. Nevertheless, the model can readily be reworked using more general processes for both micro and macro components, but the basic point remains unchanged; in deed, Pischke estimates both micro and macro income processes that are mutually consistent, and attempts to reconcile the implications with the behavior of aggregate consumption. He finds that individuals have very little incentive to obtain aggregate data; the amount that they would pay for it, calculated as the cash equivalent of the increase in *life-time* welfare from better planning, is less than 25 cents. Given that people never do learn the aggregate shock, Pischke's best fitting model, like the simple example, tends to predict too much smoothness for consumption, not too little.

Note the importance of the fact that people *never* learn the aggregate shock. In Goodfriend's model, the shock becomes known with a one period lag, so that changes in aggregate consumption, although not orthogonal to variable dated one period ago, which are effectively 'news' this period, will be orthogonal to variables dated two periods back or earlier. In consequence, lagged learning cannot explain Campbell and Mankiw's (1991) findings that changes in consumption are correlated with changes in income that were predictable on information available two periods earlier. But given the trivial benefits of learning about the aggregate, and the non-trivial costs, it seems quite plausible that individuals choose to remain in ignorance indefinitely. Note too that if Pischke's model is correct, it should in principle be possible to reject the Euler equation on the micro data, because the consumption change is not orthogonal to previously known information, namely last period's aggregate shock, although it is orthogonal to what the agent actually knew. However, such a failure would not be rejected using the sort of tests discussed in Section 1, since long time series data on individuals would be required to test whether individual consumption changes are orthogonal to aggregate shocks that are identical for everyone in the cross-section.

It is also notable that aggregation under imperfect information has an effect that is closely analogous to the effect of habit formation. Aggregate shocks are more persistent than micro shocks, but are only imperfectly perceived, so that aggregate consumption responds only slowly to aggregate income. Habits exert a direct drag on consumption change at the individual level; consumption variability is much more painful than it is in the absence of habits, and an overenthusiastic response to good luck will only give a hostage to future misfortune.

Citations

Abowd, John, and David A. Card, 1989, 'On the covariance structure of earnings and hours changes, *Econometrica*, **57**, 411–445.

Altonji, Joseph, and Aloysius Siow, 1987, 'Testing the response of consumption to income changes with (noisy) panel data,' *Quarterly Journal of Economics*, **102**, 293–328.

Browning, Martin J., Angus S. Deaton, and Margaret Irish, 1985, 'A profitable approach to labor supply and commodity demands over the life-cycle,' *Econometrica*, **53**, 503–544.

Campbell, John Y., 1987, 'Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis,' *Econometrica*, **55**, 1249–1273.

Campbell, John Y., and N. Gregory Mankiw, 1991, 'The response of consumption to income: a cross-country investigation,' *European Economic Review*, **35**, 715–721.

Carroll, Christopher D., and Lawrence H. Summers, 1991, 'Consumption growth parallels income growth: some new evidence,' pp 305–343 in B. Douglas Bernheim and John B. Shoven, editors, *National Saving and Economic Performance*, Chicago. Chicago University Press for NBER.

Chamberlain, Gary, 1984, 'Panel data,' in Zvi Griliches and Michael D. Intriligator, editors, *Handbook of Econometrics*, **2**, North-Holland. Amsterdam, 1247–1318.

Clarida, Richard H., 1991, 'Aggregate stochastic implications of the life-cycle hypothesis,' *Quarterly Journal of Economics*, **106**, 851–867.

Deaton, Angus S., 1985, 'Panel data from a time-series of cross-sections,' *Journal of Econometrics*, **30**, 109–126.

Deaton, Angus S., 1991a, 'Saving and liquidity constraints,' *Econometrica*, **59**, 1221–1248.

Deaton, Angus S., 1991b, 'Household saving in LDC's: credit markets, insurance, and welfare,' Research Program in Development Studies, Center of International Studies, Princeton University, processed.

Dow, J. Christopher R., 1965, *The management of the British economy, 1945–60*, Cambridge. Cambridge University Press.

Flavin, Marjorie, 1981, 'The adjustment of consumption to changing expectations about future income,' *Journal of Political Economy*, **89**, 974–1009.

Flavin, Marjorie, 1991, 'The joint consumption/asset demand decision: a case study in robust estimation,' NBER Working paper No 3802.

Friedman, Milton, 1957, *A theory of the consumption function*, Princeton. Princeton University Press.

Galí, Jordi, 1990, 'Finite horizons, life cycle savings and time series evidence on consumption,' *Journal of Monetary Economics*, **26**, 433–452.

Goodfriend, Marvin, 1987, 'Information aggregation bias: the case of consumption,' Federal Reserve Bank of Richmond, processed.

Grossman, Sanford J., and Robert J. Shiller, 1982, 'Consumption correlatednes and risk measurement in economies with nontraded assets and heterogeneous information,' *Journal of Financial Economics*, **10**, 195–210.

Hall, Robert E., and Frederic S. Mishkin, 1982, 'The sensitivity of consumption to transitory income: estimates from panel data on households,' *Econometrica*, **50**, 461–481.

Hayashi, Fumio, 1982, 'The permanent income hypothesis: estimation and testing by instrumental variables,' *Journal of Political Economy*, **90**, 895–916.

Hayashi, Fumio, 1985a, 'The effect of liquidity constraints on consumption: a cross-sectional analysis,' *Quarterly Journal of Economics*, **100**, 183–206.

Hayashi, Fumio, 1985b, 'The permanent income hypothesis and consumption durability: analysis based on Japanese panel data,' *Quarterly Journal of Economics*, **100**, 1083–1113.

Hotz, V. Joseph, Finn E. Kydland, and Guilherme L. Sedlacek, 1988, 'Intertemporal preferences and labor supply,' *Econometrica*, **56**, 335–360.

Jappelli, Tullio, 1990, 'Who is credit-constrained in the US economy,' *Quarterly Journal of Economics*, **105**, 219–234.

MaCurdy, Thomas E., 1982, 'The use of time-series processes to model the error structure of earnings in longitudinal data analysis,' *Journal of Econometrics*, **18**, 83–114.

Mankiw, N. Gregory, 1982, 'Hall's consumption hypothesis and durable goods,' *Journal of Monetary Economics*, **10**, 417–425.

Mariger, Randall P., and Kathryn Shaw, 1990, 'Unanticipated aggregate disturbances and tests of the life-cycle model using panel data,' Board of Governors of the Federal Reserve and Carnegie Mellon University, processed.

Mork, Knut A. and V. Kerry Smith, 1989, 'Testing the life-cycle hypothesis with a Norwegian household panel,' *Journal of Business and Economic Statistics*, **7**, 287–296.

Mroz, Thomas A., 1987, 'The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions,' *Econometrica*, **55**, 765–800.

Pischke, Jörn-Steffen, 1991, 'Individual income, incomplete information, and aggregate consumption,' Industrial Relations Section working paper no. 289, Princeton University, processed.

Runkle, David E., 1991, 'Liquidity constraints and the permanent-income hypothesis,' *Journal of Monetary Economics*, 27, 73–98.

Stone, J. Richard N., 1966, Spending and saving in relation to income and wealth, *L'Industria*, No 4, 471–499.

Stone, J. Richard N., 1973, 'Personal spending and saving in postwar Britain,' in H. Bos, editor, *Economic structure and development: essays in honor of Jan Tinbergen*, Amsterdam. North-Holland.

Visaria, Pravin, 1980, *Poverty and living standards in Asia*, The Wold Bank, Living Standards Measurement Study.

Wilcox, David W., 1987, 'Income tax refunds and the timing of consumption expenditures,' Board of Governors, Federal Reserve System, Washington, D.C., processed. (May.)

Wilcox, David W., 1989, 'Social security benefits, consumption expenditure, and the life cycle hypothesis,' *Journal of Political Economy*, 97, 288–304.

Zeldes, Stephen P., 1989, 'Consumption and liquidity constraints: an empirical investigation,' *Journal of Political Economy*, 97, 305–346.

RECENTLY PUBLISHED «TEMI» (*)

n. 140 — *The Experience with Economic Policy Coordination: the Tripolar and the European Dimensions,* by G. GOMEL - F. SACCOMANNI - S. VONA (luglio 1990).

n. 141 — *The Short-Term Behavior of Interest Rates: Did the Founding of the Fed Really Matter?,* by P. ANGELINI (ottobre 1990).

n. 142 — *Evoluzione e performance dei fondi comuni mobiliari italiani,* di F. PANETTA - E. ZAUTZIK (ottobre 1990).

n. 143 — *L'imputazione dei dati mancanti nelle indagini campionarie: un'applicazione delle tecniche di regressione,* di F. TRIMARCHI (novembre 1990).

n. 144 — *On the Measurement of Intra-Industry Trade: Some Further Thoughts,* by S. VONA ✝ (dicembre 1990).

n. 145 — *Exchange Rate Variability and Trade: Why is it so Difficult to Find Any Empirical Relationship?,* by L. BINI SMAGHI (dicembre 1990).

n. 146 — *La scelta del meccanismo di collocamento dei titoli di Stato: analisi teorica e valutazione dell'esperienza italiana,* di L. BUTTIGLIONE - A. PRATI (gennaio 1991).

n. 147 — *Diversification and Performance,* by M. BIANCO (gennaio 1991).

n. 148 — *Exchange Rate and Pricing Strategies in a Model of International Duopoly,* by P. CASELLI (gennaio 1991).

n. 149 — *Concorrenza e redditività nell'industria bancaria: un confronto internazionale,* di V. CONTI (febbraio 1991).

n. 150 — *Economie di scala e di diversificazione nel sistema bancario italiano,* di C. CONIGLIANI - R. DE BONIS - G. MOTTA - G. PARIGI (febbraio 1991).

n. 151 — *Politiche di offerta e riallocazione del credito bancario negli anni ottanta,* di C. GIANNINI - L. PAPI - A. PRATI (febbraio 1991).

n. 152 — *Stime regionali con pochi dati: analisi e simulazioni di stimatori alternativi per investimenti, occupazione e fatturato delle imprese manifatturiere,* di R. CESARI - L. F. SIGNORINI (marzo 1991).

n. 153 — *Dinamica retributiva e differenziali salariali,* di A. GAVOSTO - P. SESTITO (luglio 1991).

n. 154 — *Interessi reali, sistema impositivo ed effetto Sylos Labini,* di P. VAGLIASINDI (luglio 1991).

n. 155 — *Trasformazione delle scadenze e margine d'interesse degli istituti di credito mobiliare,* di P. SABBATINI (luglio 1991).

n. 156 — *Gli effetti della quotazione internazionale: il caso delle azioni italiane a Londra,* di F. PANETTA (agosto 1991).

n. 157 — *Grandi e piccole imprese negli anni ottanta: la ristrutturazione dell'industria in un'analisi di dati di bilancio,* di L. F. SIGNORINI (agosto 1991).

n. 158 — *Demand and Supply Shocks in Industrial Output,* di A. GAVOSTO - G. PELLEGRINI (novembre 1991).

n. 159 — *I* futures *e le opzioni sui titoli di Stato. Un'analisi del mercato e delle prospettive in Italia,* di A. SCALIA - L. TORNETTA (novembre 1991).

n. 160 — *Earnings Uncertainty and Precautionary Saving,* by L. GUISO - T. JAPPELLI - D. TERLIZZESE (febbraio 1992).

n. 161 — *Migrazioni in Europa: andamenti, prospettive, indicazioni di politica economica,* di G. GOMEL - S. REBECCHINI (febbraio 1992).

n. 162 — *Monetary Aggregates and Monetary Policy Coordination on the Way to Economic and Monetary Union: the Role of Cross-Border Deposits,* by P. GIUCCA - A. LEVY (febbraio 1992).

n. 163 — *Cross-Border Deposits and Monetary Aggregates in the Transition to EMU,* by I. ANGELONI - C. COTTARELLI - A. LEVY (marzo 1992).

n. 164 — *Young Households' Saving and the Life Cycle of Opportunities. Evidence from Japan and Italy,* by A. ANDO - L. GUISO - D. TERLIZZESE (marzo 1992).

n. 165 — *Bequests and Saving for Retirement. What Impels the Accumulation of Wealth?,* by F. BARCA - L. CANNARI - L. GUISO (marzo 1992).

(*) Requests for copies should be sent to:
Banca d'Italia - Servizio Studi - Divisione Biblioteca e Pubblicazioni - Via Nazionale, 91 - 00184 Rome.

BANCA D'ITALIA - CENTRO STAMPA