Working Paper Series

Information-Aggregation Bias

WP 91-06

Marvin Goodfriend
Federal Reserve Bank of Richmond

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Working Paper 91-6

INFORMATION-AGGREGATION BIAS

Marvin Goodfriend*

June 1991

Abstract

Aggregation in the presence of data processing lags distorts the information content of data, violating orthogonality restrictions that hold at the individual level. Though the phenomenon is general, it is illustrated here for the life cycle-permanent income model. Cross-section and pooled-panel data induce information-aggregation bias akin to that in aggregate time series. Calculations show that information-aggregation can seriously bias tests of the life cycle model on aggregate time series, cross-section, and pooled-panel data. (JEL C2, D12, D82, D91, E21)

*Research Department, Federal Reserve Bank of Richmond, Richmond, Virginia 23261. The paper has benefited from conversations with Marjorie Flavin, Robert King, Bennett McCallum, and Stephen Zeldes. It has also benefited from presentations at the Federal Reserve Bank of Richmond, Indiana University, the Konstanz Seminar on Monetary Theory and Policy, North Carolina State University, University of Pennsylvania, University of Rochester, and University of Virginia.

Introduction

Econometric method interprets economic time series as resulting from the choices of private agents interacting in well-organized markets. Individual agents are imagined to solve carefully specified dynamic and stochastic optimization problems. Decision rules yield optimal supply and demand (e.g., consumption, labor supply, and asset demands) in period t as functions of period t information on variables exogenous to the individual. The method is attractive because it yields first-order conditions that imply readily testable restrictions on time series generating processes.

Restrictions implied by theory at the individual level have been tested on aggregate data by assuming that they are invariant under aggregation. This paper shows that orthgonality restrictions implied by intertemporal optimization, rational expectations, and information processing need not hold under aggregation of randomly heterogeneous and imperfectly informed representative agents. Put more simply, aggregation creates problems for econometric estimation and evaluation of models that rely critically on the distinction between expectations and surprises at the individual level.

The life cycle-permanent income hypotheses is one such model. Tests for excess sensitivity of consumption have received substantial attention in recent years. The tests of consumption theory are based on particularly simple orthogonality restrictions. Thus, the consumption example provides a natural context within which to illustrate the information-aggregation bias that can invalidate orthogonality restrictions generated at the individual level. It must be emphasized, however, that the information-aggregation bias identified here is quite general and will interfere with the estimation and evaluation of other economic models as well.

The discussion opens in Section I with a brief review of the testable implications of the life cycle-permanent income hypothesis developed by Hall [1978]. In contrast to the discussion in Section I, which proceeds entirely at the level of a representative agent using aggregate variables, the analysis in Section II is carried out initially at the level of the individual agent and deals explicitly with the fact that individual agent income is generated by aggregate and relative components. Because aggregate income is only observable with a delay due to data processing lags, an individual's response to innovations in his own income involves signal processing. In a manner related to Lucas's [1976] Phillips curve example, aggregating individual agent decision rules yields some surprising results for the aggregate "consumption function." For example, information-aggregation bias invalidates Hall's test of the life cycle-permanent income model using aggregate data, though the test remains valid for variables, such as stock prices, that are widely known contemporaneously.

Section III shows that information-aggregation bias can be important in practice. To illustrate the point, Flavin's [1981] estimate of the excess sensitivity of consumption is reinterpreted after taking explicit account of randomly heterogeneous agent income within the framework of Section II. Section IV shows that procedures involved in testing the life cycle-permanent income model on cross-section and pooled-panel data induce information-aggregation bias akin to that in aggregate time series. Calculations show that the bias can be important for cross-section and pooled-panel data too.

I) Hall's Test of the Life Cycle-Permanent Income Model

Hall [1978] considered a conventional life cycle-permanent income model under uncertainty in which the household chooses a stochastic consumption plan to maximize the expected value of its time-additive utility function

(1) $\underset{0}{E}[\sum_{t=0}^{\infty}(1+\delta)^{-t}U(c_t)]$,

subject to the budget constraint,

(2) $\sum_{t=0}^{\infty}(1+r)^{-t}(c_t-y_t) = A_0$,

where $\underset{t}{E}$ ≡ mathematical expectation conditional on all information available in t, including c_{t-i}, y_{t-i}, and A_{t-1}, for i = 0, 1, 2, ...

δ ≡ subjective rate of time preference

r ≡ a constant real rate of interest

c_t ≡ consumption

y_t ≡ labor income

A_t ≡ assets

The first-order necessary condition for maximization of (1) subject to (2) is

(3) $_{t-1}E_1 U'(c_t) = [(1+\delta)/(1+r)]U'(c_{t-1})$.

As Hall pointed out, (3) implies that no information available in period t-1 apart from the level of consumption, c_{t-1}, helps predict consumption, c_t, in the sense of affecting the expected value of marginal utility. In particular, income or assets in period t-1 or earlier and consumption in t-2 or earlier are irrelevant once c_{t-1} is taken into account. Using (3), Hall went on to argue that the AR1 process $c_t = \lambda c_{t-1} + \xi_t$, where $_{t-1}E_1\xi_t = 0$, should be a good approximation to the stochastic behavior of consumption under the life cycle-permanent income hypothesis. The ξ_t term reflects consumption adjustments due to period t "news" on current and future income prospects.

Hall proposed and implemented tests of the life cycle-permanent income model by checking whether lagged income, lagged stock prices, or additional lags of consumption help predict consumption after one consumption lag is taken into account. On the basis of (3), he regarded evidence that variables other than consumption at one lag help predict consumption as inconsistent with the life cycle-permanent income model.

II) Randomly Heterogeneous and Imperfectly Informed Agents

Hall's test of the life cycle-permanent income model proceeds entirely at the level of a representative agent using aggregate variables. In contrast, the analysis in this section is carried out initially at the level of the individual agent and deals explicitly with the fact that individual agent income is generated by both aggregate and relative components. The goal is to derive an aggregate consumption generating process by explicitly summing decision rules of randomly heterogeneous and imperfectly informed agents.

The economy is populated by n individual agents. The i^{th} agent's income, y_t^i, is assumed to be generated as the sum of an aggregate component, $\frac{1}{n}y_t$, and a relative income component, v_t^i, such that

(4) $$y_t^i = \frac{1}{n}y_t + v_t^i,$$

where $\sum_{i=1}^{n} v_t^i = 0.$

The relative income generating process is assumed to be identical across agents. Agents differ only by their relative income innovations. They are, in effect, randomly heterogeneous representative agents. Aggregate and relative income are assumed to be uncorrelated. As in Section I, agents are assumed to be infinitely lived. One may think of aggregate and relative income as generated by ARMA processes. We may write

$$(5) \qquad y_t^i - {}_{t-1}E_1^i y_t^i = \frac{1}{n}\epsilon_t + u_t^i,$$

where ${}_{t-1}E_1^i$ = the i^{th} agent's expectation conditional on information set $I_{t-1}^i \equiv \{y_{t-1}^i, y_{t-2}^i, \ldots, y_{t-1}, y_{t-2}, \ldots\}$

$\epsilon_t \equiv$ the aggregate income innovation in period t

$u_t^i \equiv$ the relative income innovation in period t.

For what follows, it is convenient to assume that the utility function in (1) is quadratic, $U(c_t) = -[1/2](\bar{c}-c_t)^2$, so that, by (3), household consumption obeys the exact regression $c_t = \bar{c}(r-\delta)/(1+r) + [(1+\delta)/(1+r)]c_{t-1} + \xi_t$. Assume, in addition, that the real interest rate equals the rate of time preference, so household consumption follows a random walk.

If individual agent consumption obeyed the random walk version of the life cycle-permanent income model, and if aggregate income were contemporaneously observable, then using (5) we could write

$$(6) \qquad \Delta c_t^i = r\phi^A(\frac{1}{n})\epsilon_t + r\phi^R u_t^i,$$

where ϕ^A, ϕ^R ≡ the present value of a revision, at time t, in the expected path of future income in response to one unit period t innovations in the aggregate and relative components of individual agent income, respectively.

The $r\phi^A$ and $r\phi^R$ terms represent the period t revisions in permanent income upon observing the period t innovations in the aggregate and relative components of individual income. Under the random walk version of the life cycle-permanent income model, period t consumption adjusts one-for-one with these terms. Expression (18) below gives ϕ in terms of the parameters of the underlying income generating process.

With no other complications, aggregating (6) over all n agents using (5) and $\sum_{i=1}^{n} v_t^i = 0$ yields

(7) $$\Delta c_t = r\phi^A \epsilon_t.$$

Since ϵ_t is an aggregate income innovation, the disturbance in (7) is serially uncorrelated, and consumption is still a random walk in the aggregate. Moreover, ϵ_t is unpredictable on the basis of lagged income. In short, when aggregate income is contemporaneously observable, Hall's test of the life cycle-permanent income model remains valid for aggregate data.

However, aggregate income is not contemporaneously observable. Data collection and processing lags delay the observation of national income for at least one quarter. To capture this information delay, assume agents observe aggregate income with a one period lag.[1]

In this case, agents must infer as best they can the contemporaneous aggregate income innovation ϵ_t from their own contemporaneous income innovation, $y_t^i - {}_{t-1}E_1^i y_t^i$. In other words, agents must engage in signal processing in implementing their optimal consumption adjustment rule. Using equation (5), the optimal agent inference of ϵ_t conditional on observing $y_t^i - {}_{t-1}E_1^i y_t^i$ may be written

(8) $$\underset{t}{E}^{i-} \epsilon_t = n\Omega[y_t^i - {}_{t-1}E_1^i y_t^i],$$

where $\underset{t}{E}^{i-}$ ≡ the i^{th} agent's expectation conditional on information set $I_t^{i-} \equiv \{y_t^i, y_{t-1}^i, \ldots, y_{t-1}, y_{t-2}, \ldots\}$,

and where (5) and statistical independence of ϵ_t and u_t^i imply

$$(9) \qquad \Omega = \frac{(\frac{1}{n})^2\sigma_\epsilon^2}{(\frac{1}{n})^2\sigma_\epsilon^2 + \sigma_u^2}.$$

In addition to receiving new information on individual income each period, agents receive a direct observation on the previous period's aggregate income. Any discrepancy between the previous period's actual income innovation ϵ_{t-1} and last period's prediction of ϵ_{t-1} from observing $y^i_{t-1} - {}_{t-2}E^i_{2}y^i_{t-1}$ represents new information to the individual agent. Because the new information generally leads to a revision of permanent income, it also induces an adjustment in consumption.

Decision rule (6) must be modified to take the above considerations into account. First, $\frac{1}{n}\epsilon_t$ and u^i_t must be replaced with their optimal contemporaneous inferences $\Omega(y^i_t - {}_{t-1}E^i_1 y^i_t)$ and $(1-\Omega)(y^i_t - {}_{t-1}E^i_1 y^i_t)$, respectively. Second, the decision rule must include the $\frac{1}{n}\epsilon_t$ and u^i_t inference errors made known when aggregate period t income is published in period t+1. Lagged unit ϵ and u inference errors cause permanent income to be revised by $r\phi^A(1+r)$ and $r\phi^R(1+r)$ units, respectively. Hence, equation (6) becomes

$$(10) \qquad \Delta c^i_t = r\phi^A\Omega(y^i_t - {}_{t-1}E^i_1 y^i_t) + r(1+r)\phi^A[\tfrac{1}{n}\epsilon_{t-1} - \Omega(y^i_{t-1} - {}_{t-2}E^i_2 y^i_{t-1})]$$
$$+ r\phi^R(1-\Omega)(y^i_t - {}_{t-1}E^i_1 y^i_t) + r(1+r)\phi^R[u^i_{t-1} - (1-\Omega)(y^i_{t-1} - {}_{t-2}E^i_2 y^i_{t-1})].$$

Using (5), equation (10) may be rewritten

$$(11) \qquad \Delta c^i_t = r[\phi^A\Omega + \phi^R(1-\Omega)](\tfrac{1}{n}\epsilon_t + u^i_t) + r(1+r)(\phi^A - \phi^R)(1-\Omega)\tfrac{1}{n}\epsilon_{t-1}$$
$$+ r(1+r)\Omega(\phi^R - \phi^A)u^i_{t-1}.$$

Substituting with (9) for Ω in (11), it is easy to show that $\text{Cov}[\Delta c_t^i, \Delta c_{t-1}^i]$ is zero. In other words, individual consumption is still a random walk as it was in (6). $\text{Cov}[\Delta c_t^i, \Delta y_{t-1}^i]$ is also zero. Imperfect information does not invalidate Hall's test of the life cycle-permanent income model at the individual level.

Using the fact that (5) and $\sum_{t=1}^{n} v_t^i = 0$ imply $\sum_{i=1}^{n} u_t^i = 0$, sum (11) over all n agents to derive the aggregate consumption generating process

$$(12) \qquad \Delta c_t = r[\phi^A \Omega + \phi^R (1-\Omega)]\epsilon_t + r(1+r)(1-\Omega)[\phi^A - \phi^R]\epsilon_{t-1}.$$

The model underlying the aggregate "consumption function" (12) is one in which individual agents follow the random walk version of the life cycle-permanent income model exactly. Yet, since $0<\Omega<1$, unless ϕ^A equals ϕ^R aggregate consumption is not a random walk, i.e., $\text{Cov}[\Delta c_t, \Delta c_{t-1}]$ is not zero. This is surprising because the decision rules which together determine aggregate consumption are themselves random walks. Efficient information processing guarantees that individual consumption remains a random walk with randomly heterogeneous and imperfectly informed agents. However, the portion of the aggregate income innovation contemporaneously unperceived by individuals, i.e., $(1-\Omega)\epsilon_t$, affects their consumption both contemporaneously and with a lag. Relative income aggregates out but the consecutive effect of the initially unperceived aggregate income innovation does not, so changes in aggregate consumption are autocorrelated.

Because $\text{Cov}[\Delta c_t, y_{t-1}]$ is not zero unless ϕ^A equals ϕ^R, lagged aggregate income generally helps predict aggregate consumption changes generated by (12). Since that finding would cause Hall to reject the life cycle-permanent income model even though individual agents were following a version of it exactly, his test is clearly invalid for aggregate data under imperfect information.

In this example, Hall's test is protected against information-aggregation bias by a simple lagging procedure. With a one period reporting lag, the life cycle-permanent income hypothesis implies that aggregate income lagged two or more periods should not help predict aggregate consumption. Hall's test is valid here for aggregate candidate predictor variables lagged two periods or more.[2]

In practice, data are revised over a long period of time, so even aggregate predictor variables lagged beyond their initial publication date would still induce information-aggregation bias. The problem is completely circumvented by employing as instruments variables, such as interest rates and stock prices, that are widely known contemporaneously.

III) Reinterpreting Flavin's Estimate of The Excess Sensitivity of Consumption

Flavin [1981] used Hall's insights, but went beyond his "reduced form" test to develop a simple structural econometric system in which she estimated parameters measuring excess sensitivity of consumption to income. This section demonstrates the potential size of information-aggregation bias in the context of Flavin's estimation strategy. Flavin's method may be illustrated as follows. She proceeded by adding a current and seven lagged $\beta\Delta y$ terms to an equation like (7) to capture any direct effect of income on consumption apart from the effect operating through revisions in permanent income. I illustrate her consumption generating process here by adding a current and only two lagged $\beta\Delta y$ terms as follows

$$(13) \qquad \Delta c_t = \beta_0 \Delta y_t + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + r\phi^A \epsilon_t .$$

Flavin noted that the β coefficients in (13) cannot be consistently estimated by OLS because Δy_t is correlated with the disturbance term $r\phi^A\epsilon_t$. Using the fact that agents see their own income contemporaneously, Flavin pointed out that income in period t-1 and earlier cannot be "news" in period t. On that basis, she used a univariate income autoregression to eliminate y_t in (13). Assuming, for illustrative purposes, that income is generated by an AR3 process, her substitution yields the following consumption regression

(14) $$\Delta c_t = \mu + [\beta_0(\rho_1-1)+\beta_1]y_{t-1} + [\beta_0\rho_2-\beta_1+\beta_2]y_{t-2} + [\beta_0\rho_3-\beta_2]y_{t-3} + [r\phi^A+\beta_0]\epsilon_t$$

The coefficients in the income autoregression, i.e., ρ_1, ρ_2, and ρ_3, together with the y coefficients in (14) just identify the β_0, β_1, and β_2 coefficients. Flavin argued that βs significantly different from zero should be interpreted as evidence that consumers do not behave according to the life cycle-permanent income hypothesis. In her words, finding βs to be positive would indicate an excess sensitivity of consumption to current income.

By expressing the AR3 income generating process in moving average form, (14) can be rewritten as

(15) $$\Delta c_t = \mu + \pi_1\epsilon_{t-1} + \pi_2\epsilon_{t-2} + \pi_3\epsilon_{t-3} + \omega_t$$

where
$$\pi_1 \equiv \beta_0(\rho_1-1) + \beta_1$$
$$\pi_2 \equiv \pi_1\rho_1 + \beta_0\rho_2 - \beta_1 + \beta_2$$
$$\pi_3 \equiv \pi_1(\rho_1^2+\rho_2) + (\pi_2-\pi_1\rho_1)\rho_1 + \beta_0\rho_3 - \beta_2$$
$$\omega_t \equiv (r\phi^A+\beta_0)\epsilon_t + (\epsilon \text{ terms lagged more than three periods})$$

Now suppose that equation (14) is estimated using data generated by the model from Section II in which individual consumption follows a random walk, but aggregate information is received with a lag. In this case, we recover the βs as follows. Using (12) and the fact that the ϵ_{t-1}, ϵ_{t-2}, and ϵ_{t-3} in (15) are orthogonal to each other and to ω_t, we may express the π regression coefficients as[3]

(16) $$\pi_1 = r(1+r)(1-\Omega)(\phi^A-\phi^R)$$

$$\pi_2 = 0$$

$$\pi_3 = 0.$$

The βs solve equation system (16). For example, the solutions for β_0 and β_1 are

(17) $$\beta_0 = \pi_1[\frac{1-\rho_1-\rho_2}{\rho_1+\rho_2+\rho_3-1}]$$

$$\beta_1 = \pi_1[1+(1-\rho_1)[\frac{1-\rho_1-\rho_2}{\rho_1+\rho_2+\rho_3-1}]]$$

where $\rho_1+\rho_2+\rho_3-1 < 0$ required for stationarity of the AR3 income autoregression.[4]

The model underlying the expressions for β_0 and β_1 in (17) is one in which individual agents follow the life cycle-permanent income hypothesis exactly. Yet the βs would all be zero only if π_1=0 which, since $0<\Omega<1$, would require that ϕ^A equal ϕ^R. The condition $\phi^A=\phi^R$ means that the present value of a revision in the anticipated path of future income due to an innovation must be identical for both the aggregate and relative income generating components of individual income. There is no reason to expect this. In fact, Friedman [1957] exploited cross-sectional differences in income generating processes for his most striking confirmation of the permanent income hypothesis. In short, once randomly heterogeneous agent income is taken into account in deriving the aggregate consumption function, there is no reason to expect the βs to be zero even if agents follow the life cycle-permanent income hypothesis exactly. Contrary to Flavin's claim, estimating the βs to be significantly different from zero does not necessarily constitute evidence against the life cycle-permanent income hypothesis.

Information-aggregation affects Flavin's β estimates by biasing π_1 away from zero. We can assess the potential importance of the bias by focusing on π_1. First choose a plausible set of parameter values. Since Flavin used quarterly data, if we take the annual interest rate to be 4 percent, then r=0.01. Assuming the innovation variances of the aggregate and relative components of individual income to be the same, we have Ω=0.5.[5] In addition, we need to multiply the expression for π_1 by 0.5 to account for the fact that Flavin used only nondurable consumption as the dependent variable, which averages roughly half of the total over her sample period. The result is $0.0025(\phi^A-\phi^R)$.[6]

What is a plausible range for $\phi^A-\phi^R$? Flavin, pp. 988-89, shows that for an ARMA(p,q) income generating process, ϕ can be expressed as

$$(18) \qquad \phi(\rho_j,\ j=1,\ \ldots,\ p;\ \gamma_s,\ s=1,\ \ldots,\ q) = \frac{1}{1+r} \frac{1 + \sum_{s=1}^{q}(\frac{1}{1+r})^s \gamma_s}{1 - \sum_{j=1}^{p}(\frac{1}{1+r})^j \rho_j} .$$

Consider, for example, an ARMA(1,1) process. Using (18), we can write the value of ϕ corresponding to that process as

$$(19) \qquad \phi(\rho,\gamma) = \frac{1}{1+r}(\frac{1+r+\gamma}{1+r-\rho}).$$

Expression (19) makes clear that the value of ϕ can be extremely sensitive to values of the autoregressive and moving average coefficients, ρ

and γ. If ρ=1 and γ=0, then income is a random walk and ϕ=100. But values of ρ only a bit below unity can reduce ϕ substantially. For instance, ρ=0.95 and γ=0 yield ϕ=17. Finally, differences in the moving average coefficient γ, when ρ is unity or near unity, can affect ϕ by as much as a factor of 100.

Since aggregate income is near or actually nonstationary, even small differences in persistence between relative and aggregate income generating processes can yield large values for $\phi^A-\phi^R$.[7] For example, the set of parameter values chosen above requires $\phi^A-\phi^R$ to be about 23 for information-aggregation bias alone to account for Flavin's estimated π_1 value of 0.058. If aggregate and relative income generating processes were each ARMA(1,1) with ρ=1, the required $\phi^A-\phi^R$ value could be produced by a difference between their moving average (γ) coefficients of only 0.23. Alternatively, if aggregate and relative income processes were AR1 and ρ were 0.99 for aggregate income, then a ρ value of about 0.97 for relative income would yield the required $\phi^A-\phi^R$. It seems fair to say that information-aggregation can easily yield large enough values of π_1 to be important in the context of Flavin's strategy for estimating the excess sensitivity of consumption to income.

Flavin's β estimates are vulnerable to information-aggregation bias because y_{t-1} is not a valid instrument for Δy_t in (13). The reason is that y_{t-1} coincides with the one period reporting lag. However, the lagging procedure described above applies here too. In this example, Flavin's procedure is protected against information-aggregation bias by simply lagging the set of income instruments one period beyond the publication lag, e.g., by using income from t-2 and earlier to instrument for Δy_t and Δy_{t-1}. As mentioned above, the fact that actual data are revised after their initial release means that even aggregate variables lagged beyond their initial publication date may not be free of bias. Again, widely-known contemporaneously-observable instruments such as financial prices completely circumvent the problem.

It is worth pointing out that while information-aggregation bias may partly explain the rejections of the life cycle model in aggregate data, it appears not to be the sole explanation. For example, Campbell and Mankiw [1990] perform a test that is robust to the information-aggregation problem. They regress the log change in consumption on the contemporaneous log change in income, using nominal interest rate changes dated t-2 and earlier as instruments. They find a large and statistically significant coefficient.

IV) Implications for Cross-Section and Panel Data

A number of studies have been conducted implementing Hall's test of the life cycle-permanent income model on panel data.[8] Because such procedures are based on data for individual spending units, they would appear to be immune to misinterpretations arising from information-aggregation bias. This would certainly be the case if the tests were performed on individual household data separately, since the orthogonality conditions implied by (3) have to hold for individual households following the life cycle-permanent income model. However, panel data consist of a large number of households reporting over a relatively short time period. For example, Hall and Mishkin [1982] used a panel of about 2000 households reporting annual consumption over a seven-year span. Because the time series on individual households are so short, panel data is invariably pooled prior to applying statistical procedures. The purpose of this section is to illustrate how statistical procedures associated with pooling induce an information-aggregation bias akin to that in aggregate time series.

To illustrate the point, suppose that we have individual agent data Δc_t^i and Δy_t^i generated in the environment of Section II. For such data, we know that a simple regression of Δc_t^i on Δy_{t-1}^i for the i^{th} agent,

$$(20) \qquad \Delta c_t^i = a_1^i + a_2^i \Delta y_{t-1}^i + e_t^i, \qquad\qquad i = 1, 2, \ldots, n;$$

where $\quad e_t^i$ = the disturbance term,

yields $a_2^i=0$. We also know that a single regression pooling all the household data would yield a coefficient on Δy_{t-1}^i of zero. Assuming that the number of time series observations is large, consistency in the pooled regression requires only that $Cov[\Delta y_{t-1}^i, e_t^i]$ is zero, a condition guaranteed by individual behavior in accord with the life cycle-permanent income hypothesis.

The number of time series observations is, however, often quite small in panel data. For a short panel, a consistent estimate of the a_2 coefficient in the pooled regression requires the stronger condition that the e_t^i disturbances be uncorrelated. See Chamberlain [1984]. The life cycle-permanent income hypothesis guarantees that e_t^i is serially uncorrelated. But because they capture the adjustment in consumption due to news about future income prospects, which includes common components, the e_t^i disturbances are correlated across agents in any given period. If, for example, aggregate income were unexpectedly high this period, then all agents would tend to have higher than expected consumption this period.

Most studies using pooled data purge individual household consumption expectational errors of their common component to yield consistent parameter estimates and standard errors. One technique for doing so, as in Zeldes [1989], is to use wave dummies to capture the aggregate component of expectation errors. For example, one could regress Δc_t^i for each agent in the panel on the aggregate Δc_t and use only the residual, $\Delta \tilde{c}_t^i = \Delta c_t^i - E[\Delta c_t^i | \Delta c_t]$, in the pooled regression. For data generated in the environment of Section II, we would have

$$(21) \qquad \Delta\tilde{c}_t^i = r[\phi^A\Omega+\phi^R(1-\Omega)]u_t^i + r(1+r)\Omega(\phi^R-\phi^A)u_{t-1}^i.$$

The absence of ϵ terms in (21) indicates that the procedure does purge individual consumption data of its common component. However, if $\phi^A \neq \phi^R$, then neither $\mathrm{Cov}[\Delta\tilde{c}_t^i, \Delta\tilde{c}_{t-1}^i]$ nor $\mathrm{Cov}[\Delta\tilde{c}_t^i, \Delta\tilde{y}_{t-1}^i]$ would be zero. Therefore, an econometrician using the transformed series, $\Delta\tilde{c}_t^i$, would not find consumption to be a random walk. Nor would he find the a_2 coefficient to be zero, even though individual agents were following the random walk version of the life cycle-permanent income model. Note the similarity between (21) and (12). Aggregation eliminates the u terms in (12). Purging individual consumption changes of their common component eliminates the ϵ terms in (21). Both data transformations make individual agents appear inefficient relative to life cycle-permanent income behavior by implicitly imputing to them too much contemporaneous information.

The transformation would be admissible only if it were unimportant for individual agents to distinguish relative and aggregate components of their own income, i.e., if $\phi^A=\phi^R$, or if aggregate income were contemporaneously observable. In general, however, there is a tension between removing residual correlation and introducing information-aggregation bias. In short, because of statistical procedures associated with pooling, tests of the life cycle-permanent income model on panel data are not immune to information-aggregation bias. There seems to be no way of avoiding the bias for the coefficient on Δy_{t-1}^i. But in this example too, the coefficients on a set of Δy^i regressors lagged at least two periods are free of information-aggregation bias. Note once more, however, that data revisions in practice mean that aggregate variables lagged beyond their initial release date are still not completely free of bias.

In Section III, we saw that information-aggregation bias could be large in the context of Flavin's strategy for estimating the excess sensitivity of consumption to income. We also saw that since Flavin estimated $\pi_1>0$, ϕ^A must exceed ϕ^R for information-aggregation bias to contribute to the excess sensitivity she found. For $\phi^A>\phi^R$, from (21) we would predict that if individual agents followed the life cycle-permanent income model, studies using pooled-panel data would find a negative correlation between the change in consumption and lagged income. In fact, studies such as Hall and Mishkin [1982], Hayashi [1985], and Zeldes [1989] all find significant negative correlation.

The above result is intriguing. One wonders to what extent information-aggregation might be responsible for it. But we will not pursue that question here. We are merely interested in gauging the potential importance of information-aggregation for pooled-panel data. One way of doing so is to ask whether information-aggregation bias could account plausibly for the magnitude of the negative correlation. Hall and Mishkin used annual data and Zeldes used data in logarithmic form. However, Hayashi's finding is for simple quarterly changes, so his is readily compared to the results discussed above for aggregate time series.

For a panel of about 2000 Japanese households, Hayashi [1985] p. 1092, reported a highly significant correlation coefficient of -0.08 between the change in food expenditure from 1981:4 to 1982:1 and the lagged change in income. By definition, the correlation coefficient is calculated using data in deviation from mean form. In other words, Hayashi reported an estimate of $\text{Cov}[\Delta\tilde{c}_t^i,\Delta\tilde{y}_{t-1}^i]/\sqrt{(\text{Var}\Delta\tilde{c}^i)(\text{Var}\Delta\tilde{y}^i)}$, where $\Delta\tilde{y}_t^i = \Delta v_t^i$. Hayashi's correlation coefficient was estimated using only a cross section. But information-aggregation causes problems even if a cross section is used in estimation because the constant term is equivalent to a single time dummy.[9]

We can use individual consumption data generated in the environment of Section II subjected to the "~" transform, to get an idea of the size of the correlation coefficient we would expect to see in practice if individual agents followed the life cycle-permanent income model exactly but aggregate income were reported with a one period lag.[10]

To do so, we use $\Delta\tilde{c}_t^i$ from (21) to write $\text{Cov}[\Delta\tilde{c}_t^i, \Delta\tilde{y}_{t-1}^i] = r(1+r)\Omega(\phi^R-\phi^A)\sigma_u^2$ and $\text{Var}\Delta\tilde{c}^i = \left[r^2[\phi^A\Omega+\phi^R(1-\Omega)]^2+[r(1+r)\Omega(\phi^R-\phi^A)]^2\right]\sigma_u^2$.[11] The ϕ^A and ϕ^R values in these formulas depend, as discussed in Section III, on the processes that generate aggregate and relative income. The formula for $\text{Var}\Delta\tilde{y}^i$ depends on the characteristics of the relative income generating process.[12]

Consider some simple cases. Suppose aggregate income were a random walk and the first-difference of relative income were MA1. Let $r=0.01$ and $\Omega=0.5$ as assumed in Section III. In this case, a relative income MA coefficient of only -0.15 yields the -0.08 correlation coefficient found by Hayashi. Alternatively, if aggregate and relative income were both trend-stationary AR1 processes with an AR coefficient of 0.99 for aggregate income, then an AR coefficient of only 0.9865 for relative income would yield -0.08. Casual evidence might suggest that the relative income innovation variance greatly exceeds that for aggregate income innovations, so that Ω may be relatively close to zero. Even so, an Ω as low as 0.05 yields a value for Hayashi's correlation coefficient of -0.09 if aggregate income is a random walk and relative income is a trend-stationary AR1 with an AR coefficient of 0.98.

Hayashi's Table IV reports an estimated ARIMA (2,1,0) relative disposable income process with first and second order AR coefficients of −0.92 and −0.35 respectively. As a final example, consider matching Hayashi's estimated process with a realistic aggregate income process. Symmetric treatment argues for using a difference-stationary representation for aggregate income. Unfortunately, Hayashi does not report an aggregate income process. But consider one for quarterly labor income estimated on post-war U.S. data reported in Campbell and Deaton (1989), p. 361. Their aggregate income process is AR1 in growth rates with an AR coefficient of 0.443.

For these aggregate and relative income generating processes, ϕ^A=179 and ϕ^R=44.[13] In this case, Ω=0.5 yields a value for the correlation coefficient of −0.35. The more realistic value of Ω=0.05 yields a correlation coefficient of −0.09.

Before concluding, I wish to return briefly to a point touched on above. A number of other explanations have been advanced to account for the correlation between the change in consumption and the lagged change in income. These include liquidity constraints, time aggregation, adjustment costs for consumer durables, or simply the fact that income news in a given quarter may be received after some consumption decisions have already been executed.[14] None of the competing explanations, however, predicts the sign of the correlation estimated on pooled-panel data to be the opposite of that estimated on aggregate time series. The information-aggregation view alone suggests an explanation for the change of sign, thereby reconciling apparently contradictory aggregate- and panel-data findings.

Conclusion

This paper introduced the idea of information-aggregation bias, illustrated the source of the problem for aggregate time series, cross-section, and panel data, and gauged its potential importance in practical work. The paper should not be construed as demonstrating that information-aggregation actually explains the rejections in the consumption literature. Information-aggregation bias may in fact be partially responsible for some rejections in the consumption literature, and important in a host of other applications as well. Evaluating the size of the bias in specific applications, however, must be left for future work. To avoid the bias, aggregate predictor variables employed in model estimation and evaluation should be lagged one period beyond their initial publication date. Potential bias associated with data revisions is completely circumvented by employing as instruments variables that are widely known contemporaneously.

FOOTNOTES

1. Throughout the paper, I assume that agents observe their own real income contemporaneously. In fact, they only observe their nominal income currently. Only when price level information becomes available can agents calculate their own real income innovations. This source of imperfect information would be sufficient, even without heterogeneous real income, to create information-aggregation bias. It has been omitted from the discussion merely to keep the examples as simple as possible.

2. Recent work on time aggregation in consumption by Hall [1988] provides another reason to exclude the first lag of income from tests of the permanent income hypothesis. Campbell and Mankiw [1990] discuss other reasons for lagging the instruments more than one period in such tests.

3. See Goldberger [1964], pp. 200-01.

4. See the material in Box and Jenkins [1976], pp. 53-4.

For any AR(p) income generating process the expressions in (17) generalize to $\beta_0 = \pi_1[\cdot]$ and $\beta_1 = \pi_1[1+(1-\rho_1)[\cdot]]$, where $[\cdot] \equiv 1-\rho_1-\rho_2-\ldots-\rho_{p-1}/\rho_1+\rho_2+\ldots+\rho_p-1$.

The interpretation of Flavin's β estimates in (17) makes sense only if the y generating process is stationary. If, for example, y is a random walk, i.e., $\Delta y_t=\epsilon_t$, then β_0 in (13) can't be estimated, though we can still estimate the β_1 and β_2 coefficients. Using equation (12), we would find $\beta_1=r(1+r)(1-\Omega)(\phi^A-\phi^R)$ and $\beta_2=0$.

5. This assumption is made for convenience. If Ω were closer to unity because the variance of aggregate-income innovations exceeded the relative-income-innovation variance, then information-aggregation bias in aggregate time series would be smaller. See Barsky, Mankiw, and Zeldes [1986], Section 2 for a discussion of the magnitude of individual uncertainty.

6. Flavin's income series was quarterly NIPA disposable personal income in the post World War II period. It was initially released with a two month lag from 1947 to 1964I and a one month lag thereafter. But a monthly estimate of personal income has been reported with a half month lag throughout the period. While the calculations in the text indicate that information-aggregation bias could still be substantial, monthly release of income data reduces the bias as calculated there. In addition, predictive content of financial variables for aggregate income reduces further the news content of lagged quarterly income.

7. Evidence that income is approximately a random walk is found in Mankiw and Shapiro [1985] and references contained therein. If income were a random walk, then as pointed out in footnote 4, we could no longer interpret Flavin's β estimates according to (17). Flavin's π_1 regression coefficient could, however, still be interpreted as $r(1+r)(1-\Omega)(\phi^A-\phi^R)$, which would also be the interpretation of her β_1 coefficient. As Mankiw and Shapiro have

pointed out, Flavin's estimate of β_0 could no longer be interpreted as an excess sensitivity to current income in this case.

Based on her estimated AR(8) income autoregression, Flavin reported a ϕ value of about 18. The estimated ρ_1 coefficient in the income autoregression is 0.96 (0.09), however, which is well within one standard error of 1. So the true ϕ could easily be near 100.

MaCurdy [1982] presents some estimates of the persistence of income-generating processes for longitudinal data on wages and earnings.

8. See, for example, Hall and Mishkin [1982]. Hayashi [1986], and Zeldes [1989].

9. Stephen Zeldes has suggested to me that the information-aggregation issue might also arise for a panel even with no time dummies included, if the number of time series observations (T) were small and a constant term were included. Intuitively, it would seem that this effect would become less important as T increases.

10. During the period covered by Hayashi's panel, Japanese quarterly national income data was initially published with a two month lag. The first revision was released with a further three month lag. Some information on disposable income based on a sample of households was also collected and released on a monthly basis.

11. We need to multiply the covariance term by .3 to account for the fact that Hayashi's correlation coefficient was calculated using food consumption, which averages roughly 30 percent of the total over his sample. Likewise, the variance term must be multiplied by $(.3)^2$. These adjustments cancel in the correlation coefficient.

12. The relevant variance formulas for the examples discussed below are given in Box and Jenkins [1976], pp. 62 and 76. For $z_t = \rho z_{t-1} + e_t + \gamma e_{t-1}$,

$\mathrm{Var} z = [\frac{1+\gamma^2+2\rho\gamma}{1-\rho^2}]\sigma_e^2$, and for $z_t = \rho_1 z_{t-1} + \rho_2 z_{t-2} + e_t$,

$$\mathrm{Var} z = \left(\frac{1-\rho_2}{1+\rho_2}\right)\frac{\sigma_e^2}{[(1-\rho_2)^2-\rho_1^2]}.$$

13. If income (y) is generated by the difference-stationary process $\Delta y_t = \rho_1 \Delta y_{t-1} + \rho_2 \Delta y_{t-2} + e_t$, then we get $\phi = \frac{\beta}{(1-\beta)(1-\rho_1\beta-\rho_2\beta^2)}$, where $\beta = \frac{1}{1+r}$. See Campbell and Deaton [1989], p. 359.

The above ϕ formula may be applied directly to Hayashi's difference-stationary process. A slightly modified formula is required to calculate the ϕ value for Campbell and Deaton's AR1 process in growth rates. Equations (8)-(12) of their paper imply that the income process $\Delta \log y_t = \rho_o + \rho_1 \Delta \log y_{t-1} + e_t$ yields the following analog to my equation (7):

$\Delta c_t = r\phi y_{t-1} e_t$, where $\phi = \dfrac{1}{r\left[1-\rho_1(\frac{1+\mu}{1+r})\right]}$, μ is the average quarterly rate of income growth, and we require $r>\mu$. Using the estimates $\mu=0.00451$ and $\rho_1=0.433$ reported by Campbell and Deaton together with $r=0.01$ we get $\phi^{A}=179$.

14. Hall [1989] is a useful survey.

REFERENCES

Barsky, Robert B., N. Gregory Mankiw, and Stephen P. Zeldes. "Ricardian Consumers with Keynesian Propensities." American Economic Review, 76 (September 1986), 676-91.

Box, George E. P., and Gwilym N. Jenkins. Time Series Analysis: Forecasting and Control (San Francisco: Holden Day, 1976).

Campbell, John and Angus Deaton. "Is Consumption Too Smooth?" Review of Economic Studies 56 (July 1989), 357-373.

Campbell, John and N. Gregory Mankiw. "Permanent Income, Current Income, and Consumption." Journal of Business and Economic Statistics 8 (July 1990), 265-79.

Chamberlain, Gary. "Panel Data." In Handbook of Econometrics: Volume 2, Zvi Griliches and Michael D. Intriligator, eds. (Amsterdam: North Holland Publishing Co., 1984).

Flavin, Marjorie A. "The Adjustment of Consumption to Changing Expectations About Future Income." Journal of Political Economy 99 (October 1981), 974-1009.

Friedman, Milton. A Theory of the Consumption Function (Princeton: Princeton University Press, 1957).

Goldberger, Arthur S. Econometric Theory (New York: John Wiley and Sons, Inc., 1964).

Hall, Robert E. "Consumption." In Modern Business Cycle Theory, Robert J. Barro, editor, Cambridge, Mass: Harvard U. Press, 1989.

________. "Intertemporal Substitution in Consumption." Journal of Political Economy 96 (April 1988), 339-57.

________. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." Journal of Political Economy, 86 (December 1978), 971-88.

________ and Frederic Mishkin. "The Sensitivity of Consumption to Transitory Income: Estimates From Panel Data on Households." Econometrica 50 (March 1982), 461-81.

Hayashi, Fumio. "The Permanent Income Hypothesis and Consumption Durability: Analysis Based on Japanese Panel Data." Quarterly Journal of Economics 100, (November 1985), 1083-1113.

Lucas, Robert E., Jr. "Econometric Policy Evaluation: A Critique." The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy, K. Brunner and A. H. Meltzer, eds. (Amsterdam: North Holland, 1976).

MaCurdy, Thomas E. "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." Journal of Econometrics 18 (January 1982), 83-114.

Mankiw, N. Gregory and Matthew D. Shapiro. "Trends, Random Walks, and Tests of the Permanent Income Hypothesis." Journal of Monetary Economics 16 (September 1985), 165-74.

Zeldes, Stephen. "Consumption and Liquidity Constraints: An Empirical Investigation." Journal of Political Economy 97 (April 1989), 305-46.