Does Consumer Sentiment Forecast Household Spending? If So, Why?

By CHRISTOPHER D. CARROLL, JEFFREY C. FUHRER, AND DAVID W. WILCOX*

In the three months following the Iraqi invasion of Kuwait, the University of Michigan's Index of Consumer Sentiment (ICS) fell an unprecedented 24.3 index points, to its lowest level since the 1981–1982 recession.[1] This collapse in household confidence became the focus of a great deal of economic commentary and, indeed, frequently was cited as an important—if not the leading—cause of the economic slowdown that ensued.

Concern was fueled by the well-known contemporaneous correlation between the ICS and the growth of household spending. Figure 1 shows quarterly averages of the index, 1978–1993, together with the quarterly growth in real personal consumption expenditures as measured in the national income accounts (Bureau of Economic Analysis). The correlation is impressive.

Of course, it is not surprising that sentiment and the growth of spending are positively correlated. This correlation may simply reflect that, when economic prospects are poor, households curtail their spending and also give gloomy responses to interviewers. Thus, the contemporaneous correlation between sentiment and spending does not refute traditional life-cycle or permanent-income models of consumption. Nor does it necessarily make the job of forecasting changes in consumption any easier. From the point of view of an economic forecaster, the questions of interest are first, whether an index of consumer sentiment has any predictive power on its own for *future* changes in consumption spending, and second, whether it contains information about future changes in consumer spending *aside* from the information contained in other available indicators.

In Section I, we present evidence that the answer to the first question is a clear yes: we find that lagged values of the ICS, taken on their own, explain about 14 percent of the variation in the growth of total real personal consumption expenditures over the post-1954 period. Further investigation shows that the answer to the second question is probably yes as well, though here the margin is narrower and the evidence more murky. The ICS contributes about 3 percent to the $\bar{R}^2$ of a simple reduced-form equation for total personal consumption expenditures in the longer of the two sample periods we examine, but nothing in the shorter sample period (though the latter result is heavily influenced by the observation for 1980:2). For the major subcategories of spending, the contribution generally ranges between 1 percent and 8 percent. Overall, we read the evidence as pointing toward at least some significant incremental explanatory power.

Therefore, we take as given for the remainder of the paper that sentiment forecasts spending, and we turn to the issue of how that statistical relationship should be interpreted. One possible interpretation is that sentiment is an independent driving factor in the economy, and that changes in

*Carroll: Division of Research and Statistics, Stop 80, Federal Reserve Board, Washington, DC 20551: Fuhrer: Research Department, Federal Reserve Bank of Boston, Boston, MA 02106: Wilcox: Division of Monetary Affairs, Stop 71, Federal Reserve Board, Washington, DC 20551. We have benefited from the research assistance of Stephen Helwig and Christopher Geczy and the comments of an anonymous referee. The views expressed in this paper are those of the authors and not of the Federal Reserve Board, the Federal Reserve Bank of Boston, or the other members of the staff of either institution.

[1]The Conference Board's Consumer Confidence Index also plunged at the same time.

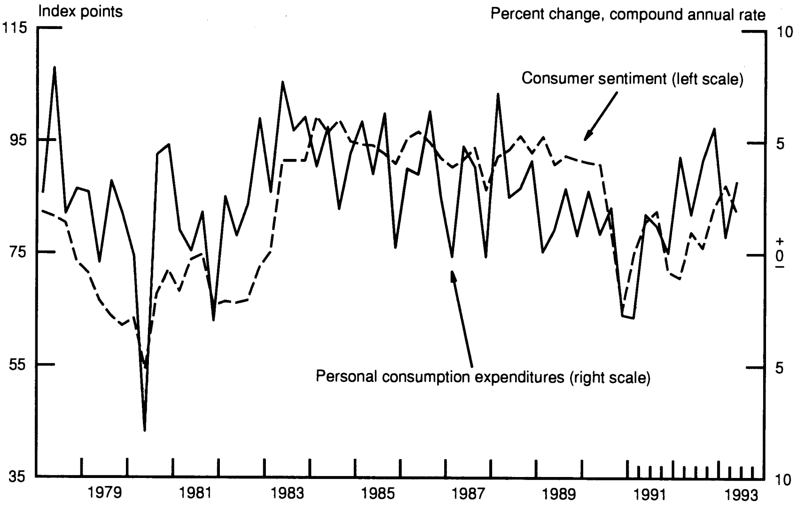

FIGURE 1. CONSUMER SENTIMENT AND THE GROWTH IN CONSUMER SPENDING

sentiment not only *forecast* changes in spending, but also *cause* them. An alternative interpretation is that sentiment forecasts spending because it reflects the overall outlook for the economy: when consumers are optimistic about the outlook for the economy, they give upbeat responses to interviewers. On average, those optimistic expectations are substantiated, and spending eventually increases as foreshadowed by sentiment.

This alternative account of the role of sentiment must involve some violation of the simplest certainty-equivalence versions of the life-cycle and permanent-income theories; otherwise, current spending would fully reflect the optimism or pessimism of households about future prospects for the economy, and sentiment would have no predictive power for future changes in spending. In Section II, we introduce such a violation using the mechanism proposed by John Y. Campbell and N. Gregory Mankiw (1989, 1990, 1991). Specifically, Campbell and Mankiw posit that some households are strict life-cyclers while others follow a "rule of thumb" and set consumption equal to income. In an economy containing those two types of consumers, sentiment might well forecast spending without being an independent driving factor: when prospects for the real economy are bright, forward-looking life-cyclers will give optimistic readings on consumer sentiment. On average, their optimism will be borne out, and income will rise. When it does, spending of rule-of-thumbers will increase. Thus, in this account, the survey responses of forward-looking households predict the spending of rule-of-thumb households.

The Campbell-Mankiw model is useful in this context because it delivers a testable implication—that sentiment should appear in the prediction equation for consumption only indirectly, in its role as a predictor of income. Our objective in Section III is to develop evidence on the validity of this explanation of the predictive content of sentiment.

In answer to the first question posed in the title, we conclude that consumer sentiment does indeed forecast future changes in household spending. Whether sentiment should be characterized as helping "a little" or "a lot" is a more difficult question. Given the amount of attention that has been paid to these indexes recently, their predictive ability seems underwhelming. From the perspective of the recent academic literature on consumption, however, their performance is impressive because it stacks up relatively well against the track record of other variables that previously have been noted to have some predictive power for personal consumption expenditures (e.g., interest rates, stock prices, the unemployment rate).

In answer to the second question, we conclude that the Campbell-Mankiw model does not provide an adequate explanation for the predictive power of sentiment for changes in household spending. Part of that predictive power appears to operate through a direct channel (or channels), rather than the indirect income channel allowed for under the rule-of-thumb account we outlined above. In the concluding section, we speculate about other possible explanations for that predictive power.

I. Some Reduced-Form Regressions

A simple way to judge the near-term predictive ability of the ICS is to examine the $\bar{R}^2$'s from regressions of the growth of various measures of household spending on lagged values of the ICS:

$$\Delta \log(C_t) = \alpha_0 + \sum_{i=1}^{N} \beta_i S_{t-i} + \varepsilon_t.$$

Note that this procedure amounts to a test of Robert E. Hall's (1978) random-walk hypothesis; if the β's are significantly different from zero, that hypothesis is rejected. Table 1 reports the results of implementing this procedure at the quarterly frequency using four lags of the ICS.[2] Over this sample period, lagged values of the ICS taken on their own explain about 14 percent of the one-quarter-ahead variation in the growth of total real personal consumption expenditures (PCE) (row 1). The probability that this explanatory power was generated merely by chance is estimated to be essentially nil (row 1, number in parentheses). As shown in rows 2–4, sentiment taken alone has the most explanatory power for one-quarter-ahead growth in PCE for goods excluding motor vehicles (17 percent of the variation explained) and the least explanatory power for PCE for motor vehicles (4 percent of the variation explained).

The second column reports results for the period since 1978, during which time the Survey Research Center has conducted its survey of consumers on a monthly basis (see footnote 2). Unfortunately, the results are rather sensitive to this change in sample period. Using the post-1978 data, we find that lagged values of the sentiment index explain only 5 percent of the variation in the growth of real PCE, 2 percent of the variation in PCE for services, and none of the variation in PCE for motor vehicles. In the case of goods excluding motor vehicles, however, the $\bar{R}^2$ is even a bit higher than it

[2] The left-hand-side variable in each regression is the log difference of the indicated category of real household spending. The starting date of the longer sample period was chosen to exclude data from the Korean War era. The starting date of the shorter sample period was chosen to coincide with the move by the University of Michigan to administer their Survey of Consumers on a monthly basis. (Prior to 1978, the survey usually was taken only once each quarter, and occasionally not even that frequently. For the period since 1978, the quarterly observation is defined to be the average of the monthly observations. For the period prior to 1978, the quarterly observation is defined to be the monthly reading if a survey was taken during that quarter, and the reading from the preceding quarter if one was not.) We ended both sample periods in 1992:3 to exclude a very dramatic fluctuation in wage and salary income in 1992:4 and 1993:1 reflecting tax-motivated shifts of income (especially bonuses and commissions) from 1993 into 1992.

We partition PCE into three categories (motor vehicles, other goods, and services) rather than the more traditional two categories (durables and nondurables plus services) to better reflect the procedures used by the Bureau of Economic Analysis to estimate consumer spending. Briefly, PCE for motor vehicles is estimated mainly from data on the unit sales of new cars and trucks; PCE for other goods is derived from the monthly retail sales data; and PCE for services is estimated from a varied collection of source data including, among others, employment in service industries, the stock of occupied housing units as estimated in the Current Population Survey, and payments for electricity and natural gas as reflected in the monthly billings of utilities. See Wilcox (1992) for further discussion. The data from the national income accounts reflect the annual revision released by the Commerce Department in September 1993. RATS code and a complete data set are available from the authors upon request.

TABLE 1—REDUCED-FORM EVIDENCE: $\bar{R}^2$'S AND INCREMENTAL $\bar{R}^2$'S FROM SIMPLE PREDICTION EQUATIONS

$$\Delta \log(C_t) = \alpha_0 + \sum_{i=1}^{4} \beta_i S_{t-i} + \gamma \mathbf{Z}_{t-1} + \varepsilon_t$$

Row	Category of real PCE	$\bar{R}^2$ 1955:1–1992:3	$\bar{R}^2$ 1978:1–1992:3	Incremental $\bar{R}^2$ 1955:1–1992:3	Incremental $\bar{R}^2$ 1978:1–1992:3
1	Total	0.14	0.05	0.03	−0.03
		(0.000)	(0.013)	(0.000)	(0.056)
2	Motor vehicles	0.04	−0.01	0.08	0.03
		(0.000)	(0.130)	(0.000)	(0.013)
3	Goods excluding motor vehicles	0.17	0.20	0.05	0.03
		(0.000)	(0.000)	(0.000)	(0.001)
4	Services	0.10	0.02	0.01	−0.07
		(0.002)	(0.030)	(0.188)	(0.969)

Notes: S_{t-i} $\{i = 1,\dots,4\}$ are lagged values of the Index of Consumer Sentiment. $\mathbf{Z}_{t-1}$ is a vector of control variables. The regressions underlying the results reported in the first two columns used only the lagged values of sentiment as explanatory variables; the regressions underlying the results reported in the third and fourth columns included the control variables. These controls included four lags of the growth in real labor income (defined as wages and salaries plus transfers minus personal contributions for social insurance). The numbers in parentheses are p values of the joint significance of the lags of sentiment. Hypothesis tests were conducted using a heteroscedasticity- and serial-correlation-robust covariance matrix (allowing serial correlation at lags up to 4).

was in the longer sample period. In each category other than motor vehicles, the coefficients on the lags of sentiment are jointly significant at the 3-percent level or better. Sensitivity to sample period notwithstanding, we interpret these results as constituting reasonably strong support for the proposition that sentiment taken alone has some predictive power for future changes in household spending.[3]

We next investigate whether the sentiment index has any predictive ability once one controls for information contained in other variables available to economic forecasters. We implement this investigation by estimating equations of the following form:

$$\Delta \log(C_t) = \alpha_0 + \sum_{i=1}^{N} \beta_i S_{t-i} + \gamma \mathbf{Z}_{t-1} + \varepsilon_t$$

where $\mathbf{Z}_t$ is a vector of other variables. Of course, the choice of *which* other variables to include in the equation is inherently somewhat arbitrary. The third and fourth columns of Table 1 present results for a minimal specification of such other variables that includes four lags of the dependent variable and four lags of the growth of real labor income, defined as wages and salaries plus transfers minus personal contributions for social insurance, all deflated by the implicit deflator for total PCE. (The use of labor income, rather than total disposable income, is motivated by our investigation in Sections II and III of the rule-of-thumb hypothesis as an explanation for the predictive power of sentiment and will be discussed in greater detail there). In these two columns, the upper entry in each cell

[3] We also estimated simple prediction equations using lags 2–5 of sentiment, rather than lags 1–4. The results were similar except for goods excluding motor vehicles, in which case the $\bar{R}^2$ fell to 8.5 percent (p value = 0.4 percent) in the longer sample period, and 4.5 percent (p value = 7.3 percent) in the shorter period.

records the *increment* to the adjusted $\bar{R}^2$ provided by the lagged values of consumer sentiment, while the lower entry (in parentheses) displays the p value from the test of the joint hypothesis that the coefficients on the four lagged values of the sentiment index equal zero.

Evidently, some—but not all—of the information in the ICS is held in common with the control variables. As the first row of the third column shows, the ICS adds only 3 percent to the explanatory power of the equation for the growth of total real PCE over the longer sample period (5 percent if the observations for 1975:2 [Social Security bonus and income-tax rebate] and 1980:2 [credit controls] are omitted). Nonetheless, the coefficients on the four lags of sentiment are estimated to be statistically significant at better than the 0.1-percent level. A similar result holds for goods excluding motor vehicles, with sentiment accounting for a smaller, but still statistically significant, proportion of the variation in the growth of spending in the presence of the control variables. In the case of services, sentiment adds only 1 percent to the $\bar{R}^2$ of the reduced-form equation, and the four lags are not jointly significant at any of the usual levels. In the motor-vehicles category, a counterintuitive result holds: the contribution to the $\bar{R}^2$ from the sentiment variables is substantially larger when the control variables are included in the regression than when they are not.

The fourth column in Table 1 repeats the experiment using the post-1978 data only, no doubt at substantial econometric risk given that 13 coefficients are being estimated in a sample of only 59 observations. If all observations are retained in the sample, the inclusion of sentiment actually slightly reduces the $\bar{R}^2$ of the prediction equation for total PCE. However, if the observation for 1980:2 is omitted (result not shown in the table), sentiment *adds* 4 percent to the $\bar{R}^2$, and the four lags are jointly significant at the 0.1-percent level. The results for services are dismal (a finding that appears to be robust to omission of various observations), but the results for motor vehicles and for goods excluding motor vehicles are quite strong. For the latter two categories of spending, the coefficients on the lags of sentiment remain jointly significant even controlling for our minimal specification of other known information. (If the observation for 1980:2 is omitted, the $\bar{R}^2$'s for motor vehicles and goods excluding motor vehicles rise to 6 percent and 7 percent, respectively.)[4]

In sum, we read these results as showing that sentiment taken on its own has considerable predictive ability for various measures of household spending. Further, sentiment likely has some (though probably not a great deal) of *incremental* predictive power relative to at least some other indicators for the growth of spending. We turn now to the interpretation of this finding.

II. Interpreting the Influence of Sentiment: The Campbell-Mankiw Framework

In a series of recent papers, Campbell and Mankiw (1989, 1990, 1991) investigate a simple modification of the pure life-cycle/permanent-income hypothesis. They assume that there are two types of consumers. One type sets spending strictly according to a standard life-cycle/permanent-income model; the other sets spending equal to current income. In the simplest version of their model, the consumption good is completely nondurable, and the decision period of consumers coincides exactly with the frequency of the data. In this case, the change in the consumption of life-cyclers is a white-noise process (equivalently, the level

[4]When we reestimate the equations underlying the results displayed in the third and fourth columns of Table 1 lagging all right-hand-side variables an additional period, the results are considerably weaker. For example, over the longer sample period, the sentiment variables subtract 0.4 percent from the $\bar{R}^2$ of the equation for total PCE rather than adding 3 percent. Over the shorter period, inclusion of sentiment subtracts from the $\bar{R}^2$ of the prediction equation for all four categories of spending. We interpret these results as illustrating the importance of retaining use of explanatory variables at the first lag, as we do in Section III.

follows a random walk):

$$\Delta C_t^{\mathrm{L}} = \varepsilon_t$$

where, as usual, ε_t represents the news received in period t about lifetime resources. Rule-of-thumb consumers set the change in their consumption equal to the change in their current income:

$$\Delta C_t^{\mathrm{R}} = \Delta Y_t^{\mathrm{R}}.$$

A crucial assumption in the Campbell-Mankiw framework is that rule-of-thumbers receive a constant proportion λ of total income. Given that assumption, aggregate consumption is given by

$$(1) \qquad \Delta C_t = \lambda \Delta Y_t + \varepsilon_t.$$

Of course, ΔY_t will be correlated with ε_t, but a consistent estimate of λ can be obtained using the standard instrumental-variables technique.[5]

Campbell and Mankiw (1989, 1990, 1991) implement a slightly more complicated version of equation (1). Many authors (e.g., Lawrence J. Christiano et al., 1991) have noted that if consumption decisions are made continuously but the data are measured as time-aggregates, the observed series on spending will follow an IMA(1,1) even if consumer behavior conforms exactly to the life-cycle model and the consumption good is completely nondurable.[6] In this case, equation (1) would be modified as follows:

$$(2) \qquad \Delta C_t = \lambda \Delta Y_t + u_t \qquad u_t \sim \mathrm{MA}(1).$$

Because u_t is serially correlated, it need not be orthogonal to variables dated $t-1$. Campbell and Mankiw (1989, 1990, 1991) address this problem by lagging their instruments an extra period (so that all instruments are dated $t-2$ or before) and by correcting their test statistics for serial correlation in the residuals. They find that, even with the instruments dated $t-2$ or before, they have enough power to reject the hypothesis that λ equals zero; their point estimates of λ center on 0.5.

A disadvantage of this estimation strategy is that it throws away the most up-to-date predictors of the change in income; from the point of view of a real-time forecaster, it would be extremely desirable to recover the use of variables dated $t-1$. This objective can be achieved by estimating equation (3):

$$(3) \qquad \Delta C_t = \lambda \Delta Y_t + \nu_t - \theta \nu_{t-1}$$

which differs from (2) only in its treatment of the error term. In equation (3), the moving-average parameter θ is estimated explicitly; as a result, one can enforce the restriction that any variable dated time $t-1$ or before should be orthogonal to ν_t even if it is not orthogonal to ν_{t-1}.[7]

We begin the next section by presenting results of estimating equation (3) using data

[5] We follow Campbell and Mankiw (1989, 1990, 1991) in applying the model to log changes in consumption and income rather than arithmetic changes.

[6] Other authors have suggested different reasons why the error term might follow an MA(1) specification. For example, Mankiw (1982) shows that the change in spending will follow an MA(1) process if the consumption good is durable. Strictly speaking, equation (2) is not consistent with the durability motivation for the MA(1) error term if the rule of thumb is understood as applying to consumption of the services of durable goods. Nonetheless, we believe that durability has much to do with the interpretation of our results. We address this issue in greater depth at the end of Section III.

[7] This estimation procedure will also be valid under certain (possibly unrealistically stringent) assumptions if the MA(1) structure of the error term is generated by measurement error in the level of consumption. If that measurement error is of the classical variety—that is, orthogonal to income and the instruments—then the resulting estimates will be consistent. There are at least two mechanisms through which this assumption could be violated. First, the measurement error in consumption could reflect the BEA's use of either income or one of our instruments as an indicator of spending. (In fact, the BEA uses information from the Bureau of Labor Statistics' labor-market surveys both to interpolate some of the detailed components of PCE for services and to construct the estimates of disposable personal income.) Second, there could be feedback from the measurement error to our instruments if, for example, prices in financial markets rise or fall in response to news of unexpectedly weak or strong estimates of economic growth.

for the four categories of spending we examined in Table 1. Such results are of interest because (i) the Campbell-Mankiw model has not (to our knowledge) been tested on data for goods and for services separately (Campbell and Mankiw [1989, 1990, 1991] used the model to explain the growth in the sum of PCE for nondurables and PCE for services); (ii) the test uses instruments dated $t-1$ and thus should have more power than the conventional implementation of the test based only on instruments dated $t-2$ or before; and (iii) our proposed explanation for the predictive power of sentiment is based on the Campbell-Mankiw model and so depends on its overall adequacy.

Lagged sentiment does not appear in equation (3) directly. Nonetheless, the Campbell-Mankiw model does not prohibit lagged sentiment from predicting current growth in consumption; it does require, however, that any predictive power of lagged sentiment for current growth in consumption should only reflect predictive power of lagged sentiment for current growth of income. In short, according to the model, lagged sentiment should enter equation (3) *at most* as an instrument for current growth of income. The alternative hypothesis is that lagged sentiment enters the prediction equation for spending directly and thus explains current growth of spending for some reason other than that it helps predict current growth of income. This alternative hypothesis is represented in equation (4):

$$(4) \quad \Delta C_t = \lambda \Delta Y_t + \sum_{i=1}^{N} \beta_i S_{t-i} + \nu_t - \theta \nu_{t-1}.$$

We test the restrictions implicit in equation (3) by estimating the more general equation (4) using the method of nonlinear instrumental variables and testing the joint significance of the β_i's.

One note on the measurement of income: previous investigators have identified Y_t with total disposable income. In our view, however, a literal interpretation of the rule-of-thumb parable suggests that Y_t should be identified with some measure of income that rule-of-thumbers would actually receive. In particular, a household that consumed all of its income would not accumulate assets, and so would not receive capital income. Consistent with this interpretation, we identified Y_t with a crude measure of labor income, constructed as wages and salaries plus transfers minus personal contributions for social insurance.[8]

III. The Campbell-Mankiw Model and Consumer Sentiment: Empirical Results

Table 2 presents 16 sets of results, reflecting the combination of four categories of household spending, two specifications of the estimating equation [equation (3), in which sentiment appears neither directly as a regressor nor indirectly as an instrument, and equation (4), in which sentiment appears both directly and indirectly], and two specifications of the instrument list. The first list of instruments comprises the regressors used in columns 3 and 4 of Table 1. The second list of instruments comprises three lags each of the dependent variable, the growth of real labor income, the change in the unemployment rate, the change in the 3-month Treasury bill rate, and the percentage change in the S&P 500 stock price index. The sample period for all results shown in Table 2 is 1955:1–1992:3.[9]

[8]Thus, our measure differs from disposable income in omitting other labor income (mainly employer contributions for pension and welfare benefit plans and directors' fees), as well as interest, dividend, rental, and proprietors' income, and in not deducting personal tax and nontax payments. The latter omission reflects the fact that the tax series sometimes is heavily influenced by fluctuations in tax payments induced by strategies probably only available to (and almost certainly only exploited by) the relatively well-off. If rule-of-thumb-type households mainly pay ordinary income tax, then the growth of their pretax income will be highly correlated with the growth of their after-tax income.

[9]If the rule-of-thumb explanation of the predictive ability of sentiment is to have any chance of holding water, then sentiment must be shown to predict the growth of real labor income. We tested this requirement by regressing the growth of real labor income on instrument lists 1 and 2, augmented with four and three lags of sentiment, respectively. Because both instrument lists include lagged spending, we estimated a total of eight regressions (four spending categories and two specifications of the remaining instruments):

TABLE 2—ESTIMATING THE CAMPBELL-MANKIW MODEL WITH AND WITHOUT LAGS OF THE INDEX OF CONSUMER SENTIMENT

$$\Delta C_t = \lambda \Delta Y_t + \sum_{i=1}^{N} \beta_i S_{t-i} + \nu_t - \theta \nu_{t-1}$$

			Without sentiment			With sentiment			
								p value on:	
Row	Category of real PCE	Instrument list	λ	θ	*p* value on overidentifying restrictions	λ	θ	Overidentifying restrictions	Joint significance of the coefficients on sentiment
1	Total	1	0.700 (0.115)	0.131 (0.081)	0.605	0.543 (0.148)	0.199 (0.072)	0.772	0.013
		2	0.734 (0.108)	0.148 (0.067)	0.332	0.612 (0.139)	0.198 (0.064)	0.335	0.002
2	Motor vehicles	1	1.786 (1.040)	0.298 (0.085)	0.736	1.217 (0.858)	0.532 (0.071)	0.673	0.000
		2	1.999 (1.022)	0.319 (0.086)	0.724	1.558 (1.986)	0.390 (0.086)	0.857	0.000
3	Goods excluding motor vehicles	1	0.801 (0.156)	0.008 (0.077)	0.330	0.584 (0.195)	0.018 (0.069)	0.314	0.049
		2	0.651 (0.136)	0.020 (0.064)	0.195	0.452 (0.160)	0.015 (0.057)	0.280	0.001
4	Services	1	0.474 (0.123)	−0.001 (0.071)	0.390	0.276 (0.126)	−0.012 (0.070)	0.674	0.014
		2	0.530 (0.092)	0.000 (0.066)	0.228	0.566 (0.104)	−0.050 (0.070)	0.212	0.719

Notes: Instrument list 1 contains a constant, ΔC_{t-i}, and ΔY_{t-i} $(i = 1,\dots,4)$; instrument list 2 contains a constant, ΔC_{t-i}, ΔY_{t-i}, ΔU_{t-i}, ΔR_{t-i}, and ΔQ_{t-i} $(i = 1,\dots,3)$, where ΔC_t is the dependent variable, Y_t is real labor income, U_t is the unemployment rate, R_t is the 3-month Treasury bill rate, and Q_t is the S&P 500 price index. The test statistic for the test of overidentifying restrictions (not shown in the table) is distributed as chi-square with degrees of freedom equal to the number of instruments less three. Figures in parentheses underneath λ and θ are standard errors. The sample period for estimation is 1955:1–1992:3.

We focus first on results obtained from estimation of equation (3), in which sentiment plays no role; these results are shown in the first three columns of the table. For total PCE, we estimate λ to be about 0.7, with a standard error of about 0.1, for either list of instruments. We estimate θ to be about 0.15; this estimate is significant at the 11-percent level when we use the first list of instruments, and at the 3-percent level when we use the second list. The test of overidentifying restrictions provides no evidence against the specification with either instrument list.

When we apply this specification to the data for the three components of PCE, the point estimates of λ differ markedly by category. For motor vehicles, we estimate λ to be about 1.8 or 2.0 (depending on the instrument list); for goods excluding motor vehicles, about 0.65 or 0.8; and for services, about 0.5. Such differences do not fit neatly into the original interpretation of λ as the fraction of income accruing to rule-of-thumb families, but they seem to correspond to the relative durability of the goods or services in the different categories, with motor vehicles (the most durable) having the highest λ and

the *p* values on the joint exclusion tests were all below 0.10, with the exception of the second instrument list when we included lagged growth of spending on services, in which case the *p* value was 0.217. Thus, we concluded that sentiment does have significant incremental predictive power for income.

with services (the least durable) having the lowest λ. Below, we provide an alternative interpretation of this parameter in which the original interpretation is appropriate for the special case in which the consumption good is completely nondurable.

The estimates of θ for goods excluding motor vehicles and for services are of trivial magnitude and statistically insignificant. For motor vehicles, we estimate θ to be about 0.3, with a standard error of slightly less than 0.1. These results are inconsistent with the hypothesis that time aggregation of a continuous-time process is the mechanism that gives rise to the moving-average structure of the error term, since if that were the right mechanism, then all three categories of spending would be expected to have moving-average parameters of about 0.25 (see Luigi Ermini, 1989). Finally, in no case do we reject the overidentifying restrictions. In fact, among these six equations, the lowest p value is 19.5 percent (goods excluding motor vehicles, using the second instrument list). In this sense, the Campbell-Mankiw model seems to provide an acceptable description of the behavior of spending in all four categories we examine.

We turn now to the tests that bear directly on the role of sentiment in the determination of spending. To execute these tests, we include lags of sentiment in the specification both as regressors and as instruments (four lags when we use the first instrument list, three lags when we use the second list), and we test whether the coefficients on the sentiment regressors are jointly significantly different from zero.

These results are presented in the remaining columns of Table 2. The results on the role of sentiment are clear: in every category except services (when we use the second instrument list), we reject the hypothesis that sentiment predicts the growth of spending only through the income channel. In most cases, these rejections are at the 2-percent level or better. As for the other aspects of the results, pairwise comparison of the first and fourth columns of the table shows that the estimates of λ are a bit smaller in most cases when sentiment is included directly than they are when sentiment is excluded, but the estimates of λ remain statistically significant except in the case of motor vehicles. The estimates of θ are somewhat larger for motor vehicles and total PCE, but still about zero for goods excluding motor vehicles and for services. In no case do we reject the overidentifying restrictions.

To investigate the robustness of our results, we estimated several variants of the basic model. First, we tried estimating equation (4) including only one lag of sentiment as an instrument and only one lag as a regressor; we found that the p values on the significance of the single sentiment regressor were higher in every case, and above the usual critical values except in the cases of total PCE with the second instrument list (5.1 percent) and goods excluding motor vehicles with the second instrument list (0.2 percent). When we included two lags of sentiment as regressors and instruments, the results resembled those we report in Table 2 much more closely: the two lags of sentiment were jointly significant at the 6-percent level or better in every category except services. We also estimated the model using a single lag of the *change* in sentiment (despite evidence derived from conventional Dickey-Fuller tests indicating that the ICS is stationary over our sample period) and obtained results similar to those that arise when we use two or more lags of the *level* of sentiment. The lagged change in sentiment was significant in every case except total PCE with the first instrument list and services with either instrument list.[10]

Second, we estimated equations (3) and (4) using the more traditional approach of lagging all the instruments (including sentiment) twice; again we found that the lags of sentiment (2–5 in the case of the first instrument list, 2–4 in the case of the second

[10]One important difference between the results obtained using one lagged change in sentiment versus two lagged levels is that the single lagged change has no incremental predictive power for the growth of labor income, whereas the two lagged levels generally do have incremental predictive power.

list) generally were jointly significantly different from zero, though the p values were more often than not a bit higher than those we report in Table 2. Finally, we reestimated the models omitting the observations for 1975:2 and 1980:2; on the whole, the results were little changed from those we report in Table 2, and the p values on the joint significance of the lags of sentiment, if anything, tended to be a bit lower.

We close this section by addressing the interpretation of our estimates of λ. Clearly, the original interpretation of λ as representing the fraction of aggregate income accruing to rule-of-thumb consumers cannot be maintained in the face of estimates that exceed 1. A natural reinterpretation can be derived, however, by supposing that rule-of-thumbers move their *consumption*, as distinct from their *outlays*, in line with contemporaneous changes in income.[11] In particular, suppose that the stock of durable goods that rule-of-thumbers wish to hold is a linear function of their income,

$$K_t^{\mathrm{R}} = \alpha Y_t^{\mathrm{R}}$$

and that durable goods accumulate according to

$$K_t = (1-\delta)K_{t-1} + X_t$$

where δ is the rate of depreciation and X_t is the rate of spending on the durable good. Then we can express the spending of rule-of-thumbers on durable goods as the following function of their income (after taking a first-order Taylor expansion):

$$(5)\quad \Delta\log\left(X_t^{\mathrm{R}}\right) = \left[\frac{1-(1-\delta)L}{\delta}\right]\Delta\log\left(Y_t^{\mathrm{R}}\right).$$

Aggregate spending on durable goods will be given (approximately) by

$$(6)\quad \Delta\log(X_t) = \lambda\left[\frac{1-(1-\delta)L}{\delta}\right]\Delta\log(Y_t) + u_t$$

$$u_t \sim \mathrm{MA}(1).$$

Equation (6) predicts that the overall coefficient on contemporaneous changes in income is an increasing function of the durability of the consumption good (a decreasing function of δ), consistent with the results shown in Table 2.[12]

IV. Conclusion

The evidence we have presented suggests that lagged consumer sentiment has some explanatory power for current changes in household spending. What sorts of explanations for that power are admissible? We ruled out the pure life-cycle/permanent-income model immediately, on the grounds that that model would admit of a *contemporaneous* correlation between sentiment and spending, but not one in which sentiment precedes spending.

We then proposed and tested a second possible explanation, based on Campbell and Mankiw's model in which some households spend according to a simple rule of thumb while the others are strict life-cyclers. In this model, lagged sentiment predicts current consumption growth only because it predicts current income growth. However, we found that in most cases we could reject the hypothesis that lagged sentiment affects consumption growth only through such an

[11] We are grateful to N. Gregory Mankiw for suggesting this alternative interpretation.

[12] Strictly speaking, equation (6) is not consistent with the specification we estimated because it introduces one lag of the growth of income as an additional regressor. Limited experimentation on our part suggested that the data have no interest in such a lagged term. We interpret this finding as consistent with those of previous authors, notably Mankiw (1982), that the predictions of this genre of models for the serial-correlation properties of spending on durable goods are very far off. Nonetheless, these models do seem to make realistic predictions concerning the income elasticities of spending on durables.

income channel. We have not formally tested other leading models of consumption, but we believe that at least the simplest versions of other models may also have some difficulty explaining our results.

If consumer sentiment is, in part, a measure of uncertainty, one might hope that a model of precautionary saving would be consistent with our results. Carroll (1992) shows that a model in which precautionary saving plays an important role does lead to correlations between uncertainty and the growth rate of consumption: an increase in uncertainty causes the level of consumption to fall, as consumers attempt to build up their stock of assets. But in subsequent periods, while the *level* of consumption will remain lower than it would have in the absence of the shock, the *growth rate* of consumption will be higher because the urgency of additional saving will wane as the stock of assets grows. Consumption growth will therefore be negatively correlated with contemporaneous uncertainty, but positively correlated with lagged uncertainty. If the consumer-sentiment index is high when uncertainty is low, this model would imply that lagged sentiment should be negatively associated with consumption growth. Instead, lagged sentiment seems to be positively correlated with consumption growth.

Another departure from the standard life-cycle/permanent-income model that has received considerable attention recently is habit formation. Karen E. Dynan (1993) shows that a simple model of habit formation implies that lagged consumption growth should have predictive power for current consumption growth. Because lagged sentiment is correlated with lagged consumption growth, it might be possible to explain a correlation of lagged sentiment with current consumption growth as arising from the correlation of lagged sentiment with lagged consumption growth. Unfortunately, such a hypothesis cannot explain our results, because lagged consumption growth was always included in our instrument sets; given that lagged consumption growth should be a sufficient statistic for expected current consumption growth, sentiment should have had no incremental explanatory power.

Thus, neither a simple model of precautionary saving nor a simple model of habit formation appears to be capable of explaining our results. However, there is some reason for thinking that a model that incorporates both habit formation and precautionary saving motives might be consistent with the pattern of facts we encounter. An increase in uncertainty in such a model would cause the desired level of consumption to be lower, but habit formation would prevent consumption from adjusting downward instantly and fully. In contrast with the simple precautionary-saving model, consumption might fall for an extended period before beginning to rise again. Furthermore, in contrast with the habit-formation model, it is no longer clear that all relevant information about the expected current growth rate of consumption should be contained in the lagged growth rate of consumption. In particular, lagged sentiment might provide incremental information about current consumption growth.[13]

To our knowledge, no formal work has been done on the short-term dynamic properties of a model that incorporates both habit formation and precautionary saving motives, so our suggestion that such a model might explain our results remains highly speculative. Nonetheless, such a model strikes us as a worthy candidate for future research in this area.

REFERENCES

Bureau of Economic Analysis. *Survey of current business.* Washington, DC: U.S. Department of Commerce, various issues.

Campbell, John Y. and Mankiw, N. Gregory. "Consumption, Income, and Interest

[13]Carroll and David N. Weil (1993) also suggest that a model incorporating both habit formation and precautionary saving might explain their empirical results on the low-frequency correlations between saving and growth in both household and aggregate data, and George M. Constantinides (1990) argues that such a model may provide an explanation for the equity-premium puzzle of Rajnish Mehra and Edward C. Prescott (1985).

Rates: Reinterpreting the Time Series Evidence," in Olivier J. Blanchard and Stanley Fischer, eds., *NBER macroeconomics annual 1989*. Cambridge, MA: MIT Press, 1989, pp. 185–216.

______. "Permanent Income, Current Income, and Consumption." *Journal of Business and Economic Statistics*, July 1990, *8*(3), pp. 265–79.

______. "The Response of Consumption to Income: A Cross-Country Investigation." *European Economic Review*, May 1991, *35*(4), pp. 723–67.

Carroll, Christopher D. "The Buffer Stock Theory of Saving: Some Macroeconomic Evidence." *Brookings Papers on Economic Activity*, 1992, (2), pp. 61–155.

Carroll, Christopher D. and Weil, David N. "Saving and Growth: A Reinterpretation." *Carnegie-Rochester Conference Series on Public Policy*, March 1993, *40*, pp. 133–92.

Christiano, Lawrence J.; Eichenbaum, Martin and Marshall, David. "The Permanent Income Hypothesis Revisited." *Econometrica*, March 1991, *59*(2), pp. 397–423.

Conference Board. *Consumer confidence survey*. New York: Conference Board, various issues.

Constantinides, George M. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, June 1990, *98*(3), pp. 519–43.

Dynan, Karen E. "Habit Formation in Consumer Preferences: Evidence from Panel Data." Board of Governors of the Federal Reserve System Economic Activity Working Paper No. 143, October 1993.

Ermini, Luigi. "Some New Evidence on the Timing of Consumption Decisions and on Their Generating Process." *Review of Economics and Statistics*, November 1989, *71*(4), pp. 643–50.

Hall, Robert E. "Stochastic Implications of the Life Cycle–Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy*. December 1978, *86*(6), pp. 971–87.

Mankiw, N. Gregory. "Hall's Consumption Hypothesis and Durable Goods." *Journal of Monetary Economics*, November 1982, *10*(3), pp. 417–26.

Mehra, Rajnish and Prescott, Edward C. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, March 1985, *15*(2), pp. 145–61.

University of Michigan. *Surveys of consumers*. Ann Arbor: University of Michigan Survey Research Center, various issues.

Wilcox, David W. "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work." *American Economic Review*, September 1992, *82*(4), pp. 922–41.