Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race

Jason Bram and Sydney Ludvigson

The effect of consumer attitudes on economic activity is a subject of great interest to both policymakers and economic forecasters. Household sentiment has been cited as one of the leading causes of the 1990-91 recession,[1] and recent levels of confidence indexes have helped fuel speculation that the economy may be headed for a period of overheating. Unexpected shifts in consumer confidence have also been used to explain swings in financial markets.

Two surveys of consumer attitudes—the University of Michigan Index of Consumer Sentiment and the Conference Board Consumer Confidence Index—are widely tracked by policymakers, financial analysts, and journalists. Despite the popularity of these indexes, there is little consensus about their ability to collect information on consumer spending that is not already captured by economic fundamentals. Also uncertain is whether one survey is more informative than the other.

In response to the widespread belief that consumers' opinions and expectations influence the direction of the economy, a growing number of studies have set out to analyze the relationship between consumer attitudes and economic variables. To date, academic research has focused exclusively on the predictive power of the University of Michigan's Index of Consumer Sentiment, most likely because of its longer time series.[2] Although these studies generally do not find a significant relationship between consumer attitudes and future real economic activity, results have varied with the economic outcomes being forecast and with the indicators included as controls.[3]

The inconclusive results of the existing research on consumer attitudes leave two important questions unanswered: Does consumer sentiment provide economically meaningful information about future consumer spending beyond that already contained in other economic indicators? Is one attitudinal measure more informative than another?

Jason Bram and Sydney Ludvigson are economists at the Federal Reserve Bank of New York.

This study is the first formal investigation of consumer attitudes that compares the forecasting power of the University of Michigan's Index of Consumer Sentiment and the Conference Board's Consumer Confidence Index. We begin with a background analysis of structural differences between the Michigan and Conference Board indexes. We then undertake a formal statistical comparison of the predictive power exhibited by each overall survey and its component questions for several categories of consumer spending growth.

The University of Michigan's Consumer Sentiment Index and the Conference Board's Consumer Confidence Index are the most widely followed measures of U.S. consumer confidence.

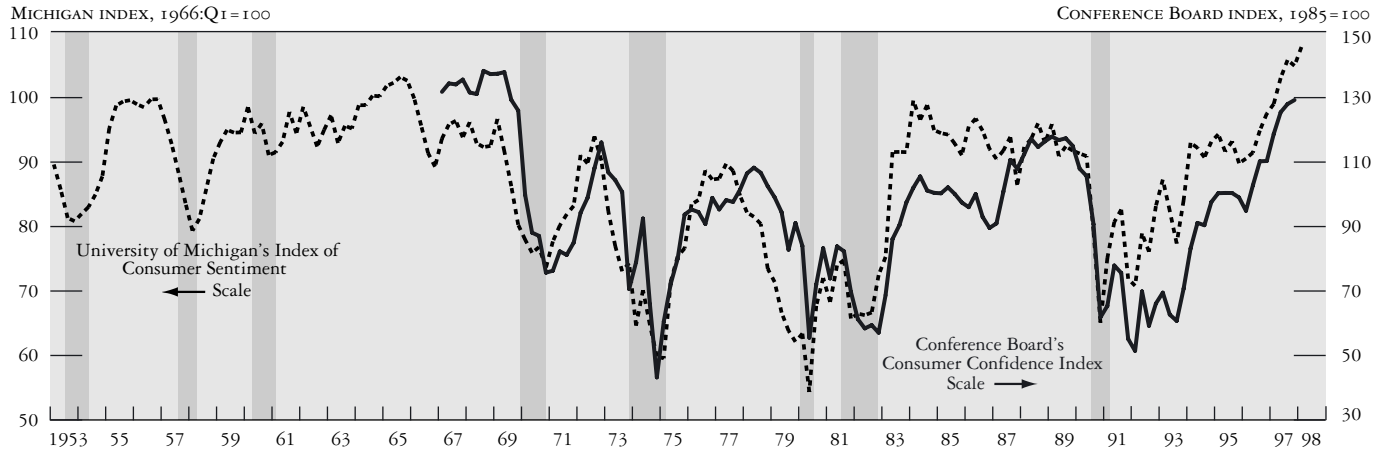
Our empirical analysis suggests that consumer sentiment can help predict future movements in consumer spending; that forecasting power, however, depends on the survey in question. Measures of consumer attitudes available from the Conference Board have both economically and statistically significant explanatory power for several spending categories—including total personal consumption expenditures; motor vehicles; services; and durables, excluding motor vehicles—even when the information contained in other economic indicators such as income, interest rates, and stock prices is known. Measures available from the University of Michigan's Survey Research Center, however, exhibit weaker forecasting power for most categories of consumer spending.[4]

A Comparison of the Michigan and Conference Board Surveys

The University of Michigan's Consumer Sentiment Index and the Conference Board's Consumer Confidence Index are the most widely followed measures of U.S. consumer confidence (Chart 1). Although the financial markets and the business community closely follow both indexes, virtually all published academic research focuses on the Michigan index—most likely because of its longer history. The Michigan index began as an annual survey in the late 1940s. In 1952, it was converted to a quarterly survey and in 1978 to a monthly survey. The Conference Board launched its index on a bimonthly basis in 1967 and expanded it to a monthly series in 1977.

Chart 1

Two Indexes of Consumer Attitudes



Sources: Conference Board; University of Michigan Survey Research Center.

Note: Shaded areas denote periods designated recessions by the National Bureau of Economic Research.

Although the two indexes broadly measure the same concept—public confidence in the economy—they are based on different sets of questions and sometimes give conflicting signals. In order to interpret movements in these two series, it is important to understand some key differences in the specific questions that are asked as well as in sample size, survey methodology, and index formulation.

Survey Questions: Present Conditions and Expectations Components

Both the Conference Board and the University of Michigan base their overall index of consumer confidence on five questions that are part of a broader survey of consumer attitudes and expectations (Box A). In addition to the overall index, both organizations report two component indexes.

Present Conditions Component

In each survey, two of the five questions ask respondents to assess present economic conditions. Michigan calls the component index based on these two questions *current conditions*, while the Conference Board uses the term *present situation*. Throughout the article, we use the generic term *present conditions* for both organizations. The present conditions questions receive a 40 percent weight in each overall index.

The Conference Board's present conditions component takes a "snapshot" approach, asking respondents to evaluate current business conditions and job availability. Because of the nature of the questions, the Conference Board's present conditions component closely tracks the nation's unemployment rate, and year-over-year changes in the index are closely correlated with payroll employment growth.

Box A: Component Questions of Consumer Confidence

Five questions make up the confidence indexes reported by the University of Michigan and the Conference Board. Each set of questions asks respondents to assess present and future economic conditions and is part of a broader monthly survey of consumer attitudes.[a]

Michigan Survey	Conference Board Survey
Present Conditions Questions	Present Conditions Questions
Q1) Do you think now is a good or bad time for people to buy major household items? [good time to buy/uncertain, depends/bad time to buy]	Q1) How would you rate present general business conditions in your area? [good/normal/bad]
Q2) Would you say that you (and your family living there) are better off or worse off financially than you were a year ago? [better/same/worse]	Q2) What would you say about available jobs in your area right now? [plentiful/not so many/hard to get]
Expectations Questions	Expectations Questions
Q3) Now turning to business conditions in the country as a whole—do you think that during the next twelve months, we'll have good times financially or bad times or what? [good times/uncertain/bad times]	Q3) Six months from now, do you think business conditions in your area will be [better/same/worse]?
Q4) Looking ahead, which would you say is more likely—that in the country as a whole we'll have continuous good times during the next five years or so or that we'll have periods of widespread unemployment or depression, or what? [good times/uncertain/bad times]	Q4) Six months from now, do you think there will be [more/same/fewer] jobs available in your area?
Q5) Now looking ahead—do you think that a year from now, you (and your family living there) will be better off financially, or worse off, or just about the same as now? [better/same/worse]	Q5) How would you guess your total family income to be six months from now? [higher/same/lower]

[a] To compare the two indexes, we reorder the questions and number them one through five. In addition, because the University of Michigan and the Conference Board use slightly different terminology for the index component based on the first two questions, we adopt the term *present conditions* for both organizations.

Michigan asks respondents to comment on the advisability of big-ticket household purchases and to assess changes in their own financial situation. Michigan's present conditions component is less closely tied to labor market conditions and its level tends to reflect recent changes in the economy rather than the level of economic activity.

These differences are reflected in the cyclical behavior of the two present conditions component indexes: Michigan's generally peaks in the early stages of economic recovery, when growth is high. By contrast, the Conference Board's generally peaks in the late stages of economic expansion, when unemployment is low and the level of economic activity is high. Not surprisingly, given the differences in the questions, the present conditions components of the two indexes are not closely correlated (Chart 2).

Expectations Component

The three questions that ask about consumers' expectations are fairly comparable in the two surveys. The Conference Board survey asks about expected changes in business conditions, job availability, and respondents' income over the next six months.[5] Michigan's poses questions on expected business conditions—both over the next year and over the next five years—and expected changes in the respondent's financial situation over the next year.[6]

Unlike the present conditions components, the expectations components in the two surveys are highly correlated with each other (Chart 3). Moreover, Michigan's present conditions and expectations components are much more closely correlated than are the Conference Board's (Appendix A).

Methodology

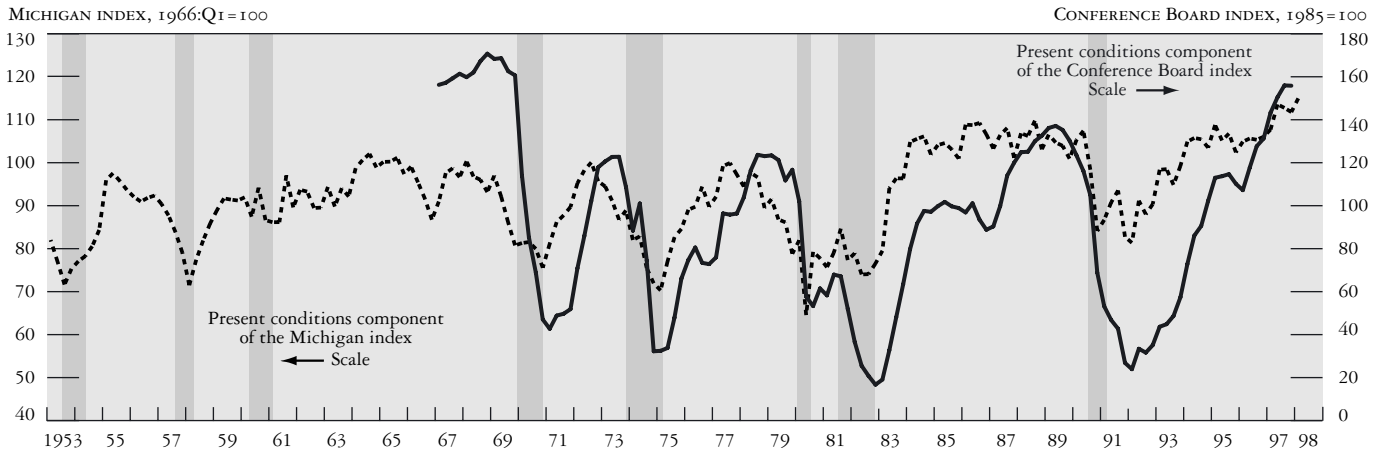
The most important methodological differences between the two surveys concern sample size, which affects sampling error and thus reliability, and index construction, which affects the range of movement in the indexes. The survey timing and release schedules also differ—a relevant consideration when conducting real-time analysis.

Michigan conducts its survey by phone throughout most of the month. Its sample size is 500; a preliminary midmonth release is based on about 250 phone interviews conducted early in the month. Final figures for the full sample are subsequently made available at the end of the month and are not subject to further revision.

The Conference Board sends out a mail survey at the end of the prior month and responses flow in throughout the survey month. The sample size is roughly 3,500 (of a total mailing of 5,000).[7] On the last Tuesday of the survey month, the Conference Board formally releases its preliminary figures

Chart 2

PRESENT CONDITIONS COMPONENT OF CONSUMER ATTITUDES

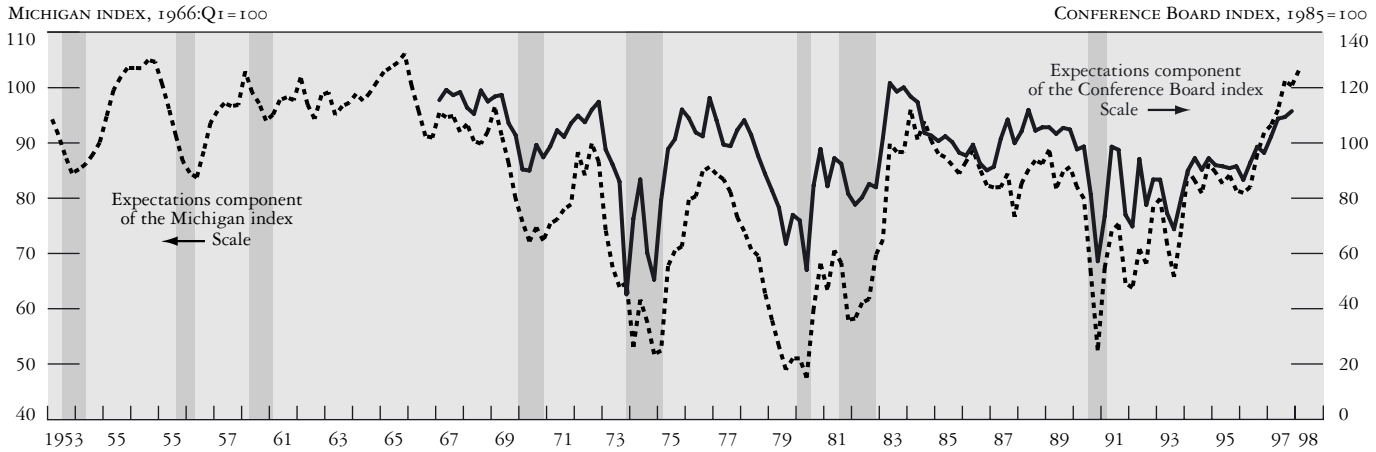


Sources: Conference Board; University of Michigan Survey Research Center.

Note: Shaded areas denote periods designated recessions by the National Bureau of Economic Research.

Chart 3

EXPECTATIONS COMPONENT OF CONSUMER ATTITUDES



Sources: Conference Board; University of Michigan Survey Research Center.

Note: Shaded areas denote periods designated recessions by the National Bureau of Economic Research.

based on about 2,500 responses. Final, revised data based on the full monthly sample are released with the next month's preliminary figures and are not subject to further revision.

The University of Michigan and the Conference Board also use different methodologies to construct their indexes from the raw response data (Box B). The main result of these methodological differences is that the Conference Board's overall index and component measures have a wider range of movement than Michigan's. In other words, identical shifts in the underlying responses tend to produce significantly larger moves in the Conference Board's indexes than in Michigan's.

A good rule of thumb is that a one-point move in Michigan's index is roughly comparable to a two-point move in the Conference Board's Index.

Interpretation of the Indexes

Although the Conference Board and Michigan indexes are highly correlated, they sometimes move independently of one another. Because of differences in survey methodology, index construction, and base year, index levels are not comparable; monthly changes must be compared on a standardized basis rather than in absolute terms. A good rule of thumb is that a one-point move in Michigan's index is roughly comparable to a two-point move in the Conference Board's index.

The indexes also differ in timeliness and reliability. One advantage of Michigan's index is that its preliminary figures are available earlier than the Conference Board's. However, because Michigan's figures are based on a much smaller sample size than the Conference Board's, they are more susceptible to measurement error. As a result, random monthly fluctuations tend to be more pronounced in Michigan's index than in the Conference Board's.[8]

Two of the most common dilemmas in relying on consumer confidence as an economic indicator are whether to focus on index level or month-to-month changes and whether to focus on the present conditions or the expectations component. For the Conference Board index, it is particularly useful to examine the present conditions and expectations components individually. The level of the present conditions component serves as a good proxy for the level of economic activity, while the expectations com-

Box B: Calculating the Two Indexes

The example below illustrates how the Conference Board and Michigan would construct a single index for one question using the same raw response data. Hypothetical figures are shown for two months along with the base-period levels against which the indexes are benchmarked.

Michigan calculates a diffusion measure by adding the difference between the positive and negative percentages to 100. Thus, the current month's value is 112 [100 + 24 – 12], and the prior month's level is 120 [100 + 30 – 10]. Next, an index is constructed by dividing the level of the diffusion measure by the base-period level of 110, and then multiplying by 100. In the example below, this calculation yields a value of 101.8 [120 + 110 – 100] for the current month, down from the prior month's level of 109.1 [120 + 110 – 100] — a drop of 7.3 points.

Example: Calculating Index Levels from Raw Response Data

	Base Period	Prior Month	Current Month
Percentage of responses			
Positive	25	30	24
Neutral	60	60	64
Negative	15	10	12
Indicator level			
Michigan diffusion measure	110.0	120.0	112.0
Michigan index	100.0	109.1	101.8
Conference Board diffusion measure	62.5	75.0	66.7
Conference Board index	100.0	120.0	106.7

Using the same raw responses, the Conference Board would calculate its diffusion measure by dividing the positive response percentage by the sum of the positive and negative response percentages. This procedure gives a value of 66.7 [$24 \div (24 + 12) \times 100$] for the current month and 75 [$30 \div (30 \times 10) \times 100$] for the prior month. Next, the index is calculated to be 106.7 [$66.7 \div 62.5 \times 100$] in the current month, down from a level of 120 [$75.0 \div 62.5 \times 100$] in the prior month—a drop of 13.3 points.

Some subtle differences in index construction are not illustrated here. First, the Conference Board converts each diffusion index to a base-year index and then averages the indexes together.[a] Michigan first averages the diffusion indexes into a composite diffusion index and then converts the results to a base-period index. Second, the Conference Board's responses are seasonally adjusted, while Michigan's are not. However, the seasonal adjustment has little effect on our results, because neither index exhibits much seasonality. Finally, because the Conference Board and Michigan use different base periods (1985 and 1966:Q1, respectively), the response patterns on which the indexes are based may differ. As a result, the index *levels* of the two surveys are not comparable.

[a] Because the Conference Board's diffusion measures are converted into base-year indexes before they are averaged arithmetically, a given question's effective weight in the index is influenced by the selection of the base year. In theory, the choice of the base year could affect the magnitude and even the direction of change in the index. (The resulting problems are similar to those associated with the old fixed-base-year GDP deflator.) In practice, however, this feature has no discernible effect on the Conference Board's index.

ponent is more closely correlated with the rate of economic growth. In Michigan's survey, both components are closely correlated and in general serve as an indicator of the pace of economic growth.

The Conference Board index, the Michigan index, and the components of each index exhibit some movement that cannot be explained by movements in other economic indicators such as income, interest rates, and lagged consumption. In the next section, we determine whether this independent movement contains information that can help predict consumer spending.

Empirical Results

We use a two-step procedure to determine the forecasting power of consumer confidence. First, we consider a baseline forecasting equation for consumption growth that does not include attitudinal survey measures. We then add consumer sentiment to the baseline equation and test which measures of consumer attitudes, if any, improve the forecasting power of the baseline equation. In estimating the confidence-augmented equation, we employ two types of statistical tests to determine whether consumer attitudes help predict future movements in consumer spending: in-sample regressions and out-of-sample regressions of consumption growth. The in-sample tests investigate the pre-

dictive power of consumer sentiment over the entire sample period; the out-of-sample procedure tests the stability of that predictive power over several subsamples of the data.

Our analysis measures the effect of consumer attitudes on five categories of household personal consumption expenditure: total expenditure; motor vehicle expenditure; expenditure on all goods, excluding motor vehicles; expenditure on services; and expenditure on durable goods, excluding motor vehicles. The data are quarterly and span the period from the first quarter of 1967 to the third quarter of 1996.[9] Definitions of the variables used in the equations appear in Appendix B.

Baseline Forecasting Equation

We specify a simple forecasting equation for consumption growth that does not include consumer confidence. This specification, or the baseline equation, takes the form

$$\Delta \ln(C_t) = \alpha_0 + \gamma Z_{t-1} + \varepsilon_t, \quad (1)$$

where C_t is real consumption spending and Z_{t-1} is a vector of control variables. In choosing economic indicators to include in Z_{t-1}, we adhere to the existing literature closely. In an earlier work, Carroll, Fuhrer, and Wilcox (1994) estimated a similar equation to test whether the Michigan index contained any incremental predictive power for future movements in consumer spending. Their baseline equation placed lagged values of the dependent variable and of labor income growth in Z. The inclusion of labor income growth as a control variable is motivated by a large and growing body of empirical work showing that consumption growth is related to lagged, or predictable, income growth (see, for example, Flavin [1981] and Campbell and Mankiw [1989]). Like Carroll, Fuhrer, and Wilcox, we include these indicators on the right-hand-side of the equation using four lags of each variable. As is typical in aggregate time series, Akaike and Schwarz tests did not indicate the need for more than four quarterly lags.

Other researchers have argued that the information contained in attitudinal indicators should be assessed relative to that contained in financial indicators. Leeper (1992) points out that consumer sentiment may have predictive power for spending because consumer surveys are made available on a more timely basis than other economic indicators such as income and consumption data. However, he goes on to argue that financial market indicators are available on an almost continuous basis and may contain much of the same information captured by consumer sentiment. Indeed, Leeper finds that consumer attitudes are only weakly correlated with variables such as unemployment and industrial production once financial indicators are included. To investigate whether consumer attitudes contain useful information for future consumer spending beyond that contained in financial indicators, we follow Leeper's suggestion and include the log first difference of the real stock price and the first difference of the three-month Treasury bill rate in our Z vector.

Our analysis measures the effect of consumer attitudes on five categories of household personal consumption expenditure: total expenditure; motor vehicle expenditure; expenditure on all goods, excluding motor vehicles; expenditure on services; and expenditure on durable goods, excluding motor vehicles.

As a robustness check for our Z specification, we substitute the unemployment rate for labor income growth. In addition, we substitute three different variables for the first difference of the three-month Treasury bill rate: the spread between the ten-year Treasury bond rate and the one-year Treasury bill rate,[10] the first difference of the one-year Treasury bill rate, and the first difference of the federal funds rate. The results, which are not reported here, indicate that these substitutions do not qualitatively alter the estimation of the baseline model. To summarize, the control variables included in Z are four lags of the

dependent variable, four lags of the growth in real labor income, four lags of the log first difference in the real stock price index as measured by the Standard and Poor's 500 index, and four lags of the first difference of the three-month Treasury bill rate.

According to our estimation, lagged values of consumption growth and the financial indicators in Z have predictive power for most categories of consumer expenditure. Table 1 presents the estimation results of the baseline model. For each category of consumption, the table presents the sum of the coefficients on the lags of each variable in Z. The sum of the coefficients on the four lags of each variable estimates the long-run effect of the variable on consumption growth. The p-values for the joint marginal significance of the lags of each variable, which appear in parentheses, give the probability that the explanatory variable can be excluded from the forecasting equation.[11] When the p-values are very low, the variables are statistically significant predictors of consumption growth.

As Table 1 shows, the long-run impact of most variables has the expected sign. Consumption growth is positively related to lagged consumption growth for most of the categories, while lagged interest rates have a small negative effect on future consumption. Interestingly, the inclusion of the consumption and interest rate variables appears to reduce the statistical significance of the income and stock market variables in forecasting consumption growth.

Table 1
BASELINE FORECAST OF CONSUMPTION GROWTH

Predicted Variable	Four Lags of Consumption	Four Lags of Income	Four Lags of Treasury Bill Rate	Four Lags of S&P 500
Total	0.83	0.04	-0.002	-0.01
	(0.000)	(0.263)	(0.006)	(0.056)
Motor vehicles	0.47	0.40	-0.024	-0.05
	(0.230)	(0.221)	(0.068)	(0.012)
Goods, excluding motor vehicles	0.88	0.04	-0.001	0.0
	(0.000)	(0.356)	(0.094)	(0.148)
Services	0.05	0.50	-0.007	-0.02
	(0.021)	(0.102)	(0.000)	(0.276)
Durables goods, excluding motor vehicles	0.80	0.16	-0.006	0.0
	(0.000)	(0.886)	(0.013)	(0.477)

Source: Authors' calculations.

Notes: The table reports the sum of the coefficients on the lags of the variable indicated; the probability that the variable can be excluded from the prediction equation appears in parentheses. Hypothesis tests were conducted using a heteroskedasticity and serial correlation robust covariance matrix. The sample covers the period from the first quarter of 1968 to the third quarter of 1996. S&P=Standard and Poor's.

ADDING CONSUMER CONFIDENCE TO THE BASELINE EQUATION

To determine whether consumer attitudes help forecast future consumer spending, we add a measure of consumer confidence to the baseline equation:

$$(2) \qquad \Delta \ln(C_t) = \alpha_0 + \sum_{i=1}^{n} \beta_i S_{t-i} + \gamma Z_{t-1} + \varepsilon_t,$$

where S is consumer confidence as measured by either the Michigan or the Conference Board index. We then replace the overall index with the expectations component as our measure of S.[12]

Our modified equation attempts to quantify the power of each index to predict future consumption expenditures. In our estimations, we report the *increment* to the adjusted R^2 that results from augmenting the baseline equation to include each of the attitudinal indicators. For example, if the increment to the adjusted R^2 from adding the four lags of S is X percent, the confidence-augmented equation predicts about X percent more of the variation in the next quarter's consumption than the baseline equation.

The first two columns of Table 2 present the results of estimating the confidence-augmented equation. The first column of Table 2 reports the results for the equation that includes the Michigan overall index (rows 1-5) and its expectations component (rows 6-10); the second column of Table 2 reports the results for the equation that includes the Conference Board overall index (rows 1-5) and its expectations component (rows 6-10). The probability that the confidence variables can be excluded from the forecasting equation appears in parentheses.[13]

Our results reveal a gap in the indexes' forecasting power for total personal consumption growth. For the Michigan survey, the lagged values of consumer sentiment do not increase the adjusted R^2 in the regression where total personal consumption growth is the dependent variable. Indeed, the inclusion of Michigan's overall index actually weakens the predictive power of

the baseline equation. We obtain similar results using the Michigan expectations component. By contrast, both the Conference Board's overall measure of consumer confidence and its measure of consumer expectations are incrementally informative about the future path of total personal consumer spending growth. Adding the last four quarters of data from the Conference Board's overall confidence index to the baseline equation predicts an additional 9 percent of the variation in the next period's consumption growth. Similarly, adding the last four quarters of data on the expectations component predicts an additional 12 percent of the variation in future consumer spending. Moreover, the Conference Board index is statistically significant at better than the 5 percent level.[14]

For motor vehicle spending, however, both overall indexes display some incremental predictive power. Lagged values of the Michigan sentiment index explain an additional 5 percent of the growth in motor vehicle spending, a relatively small amount, although the increase is statistically significant at the 6 percent level. By including lags of the Michigan expectations component, however, we increase the fraction of regression variance explained by consumer confidence to 8 percent, and the expectations variables become significant at the 5 percent level. The Conference Board measures have an equal or somewhat stronger incremental impact on growth in motor vehicle spending; the overall index, like the Michigan index, increases the adjusted R^2 by 5 percent, but the inclusion of four lags of the the Conference Board's expectations component increases the adjusted R^2 by 10 percent.[15]

For spending on services and durable goods (excluding motor vehicles), lagged values of either Michigan's overall index or its expectations component generally add little or no explanatory power to the consumption growth regressions. For services spending growth, the incremental adjusted R^2 is negative. The Michigan index does help to forecast growth in the goods (excluding motor vehicles) category. Still, even in this case, the inclusion of four lags of Michigan's overall index improves the forecasting performance of the baseline equation by just 3 percent.

Adding the last four quarters of data from the Conference Board's overall confidence index to the baseline equation predicts an additional 9 percent of the varation in the next period's consumption growth.

Table 2
FORECAST OF CONSUMPTION GROWTH, AUGMENTED BY CONSUMER CONFIDENCE INDICATORS

Real Personal Consumption Expenditures	Michigan Index	Conference Board Index	Both
	Overall Index		
Total	-0.04	0.09	0.13
Conference Board	—	(0.001)	(0.000)
Michigan	(0.715)	—	(0.040)
Motor vehicles	0.05	0.05	0.21
Conference Board	—	(0.020)	(0.000)
Michigan	(0.059)	—	(0.000)
Goods, excluding motor vehicles	0.03	0.07	0.05
Conference Board	—	(0.177)	(0.392)
Michigan	(0.000)	—	(0.934)
Services	-0.02	0.02	0.11
Conference Board	—	(0.062)	(0.001)
Michigan	(0.607)	—	(0.140)
Durables, excluding motor vehicles	0.00	0.15	0.17
Conference Board	—	(0.005)	(0.041)
Michigan	(0.257)	—	(0.780)
	Expectations Component		
Total	-0.03	0.12	0.11
Conference Board	—	(0.000)	(0.000)
Michigan	(0.557)	—	(0.645)
Motor vehicles	0.08	0.10	0.19
Conference Board	—	(0.006)	(0.000)
Michigan	(0.042)	—	(0.014)
Goods, excluding motor vehicles	0.00	-0.12	-0.02
Conference Board	—	(0.334)	(0.696)
Michigan	(0.858)	—	(0.884)
Services	-0.01	0.06	0.07
Conference Board	—	(0.018)	(0.010)
Michigan	(0.554)	—	(0.253)
Durables, excluding motor vehicles	0.03	0.06	0.02
Conference Board	—	(0.217)	(0.677)
Michigan	(0.298)	—	(0.687)

Source: Authors' calculations.

Notes: The table reports the increment to the adjusted R^2 statistic from adding four lags of the confidence measures; *p*-values for the joint marginal significance of the lags of the confidence measures appear in parentheses. Hypothesis tests were conducted using a heteroskedasticity and serial correlation robust covariance matrix. The sample covers the period from the first quarter of 1968 to the third quarter of 1996.

Lagged values of the Conference Board's overall index appear to be of value in predicting spending in durables and services. For durable goods (excluding motor vehicles), adding lags of the overall Conference Board index increases the fraction of regression variance explained by consumer confidence by 15 percent, a finding that is highly statistically significant. Moreover, the Conference Board expectations component appears strongly related to future services expenditures—the addition of that variable increases the adjusted R^2 by a statistically significant 6 percent.

For the category of motor vehicle expenditure, including both measures of consumer attitudes in the forecasting equation may be superior to the use of either index alone.

Head-to-Head Forecasting Competition

The results discussed above suggest that the Conference Board index generally serves as a better predictor of spending than the Michigan index. Despite our finding that the Michigan index has little explanatory power for categories of expenditure other than motor vehicles, it is still possible that the Conference Board index contains only a small amount of information that is independent of that contained in the Michigan index. If this were true, including both indexes in the equation simultaneously could substantially reduce the explanatory power of the Conference Board index. To examine this possibility, we estimate a "head-to-head" forecasting equation that includes both measures of consumer attitudes in the equation at the same time and takes the form

$$(3)\qquad \Delta \ln(C_t) = \alpha_0 + \sum_{i=1}^{4} \beta_i S^c_{t-i} + \sum_{i=1}^{4} \delta_i S^m_{t-i} + \gamma Z_{t-1} + \varepsilon_t ,$$

where S^c and S^m are the consumer confidence variables as measured by the Conference Board index and the Michigan index, respectively. As in equation 2, we report results when S^c and S^m are measured as each survey's overall index or its expectations component.

The results of estimating equation 3 appear in the third column of Table 2. The numbers reported for both indexes are the increment to the adjusted R^2 after both confidence measures are added to the baseline equation. The probability that the Conference Board and Michigan indexes can be excluded from equation 3 appears in parentheses. The table shows that the Conference Board variables remain statistically significant once the Michigan variables are included. Thus, the direct inclusion of both the Michigan index and the Conference Board index in the forecasting equation does not eliminate the forecasting power of the Conference Board index. Indeed, for the category of motor vehicle expenditure, including both measures of consumer attitudes in the forecasting equation may be superior to the use of either index alone. The increment to the adjusted R^2 from adding both overall sentiment measures to the motor vehicle baseline regression is 21 percent, a large increase over that obtained when the equation incorporates only one of the indexes. Moreover, both indexes remain statistically significant predictors of motor vehicle expenditure in the head-to-head specification. We discuss one possible explanation for this finding in the question-level analysis below.

As we have shown, the results in this section suggest that a gap exists in the predictive power of the two attitudinal surveys, with the Conference Board index generally outperforming the Michigan index. We now examine whether this gap can be explained by differences in the individual questionnaires.

Question-Level Analysis

The underlying questions of the Conference Board and Michigan indexes serve as mini-diffusion indexes that are similar in construction to the overall indexes.[16] We test the predictive power of each question-level index using the following equation:

$$(4)\qquad \Delta \ln(C_t) = \alpha_0 + \sum_{i=1}^{4} \beta_i Q^k_{jt-i} + \gamma Z_{t-1} + \varepsilon_t ,$$

where Q^k_j denotes question j of index k, for j=1,...5 and k=1,2.

As we would expect, no single question helps predict all categories of spending growth. Several questions, however, help predict growth in particular categories of expenditure. Table 3 reports the increment to the adjusted R^2 from adding four lags of each question to the baseline equation. As the table shows, questions two and four of the Conference Board index explain a substantial portion of the regression variance for total consumption, motor vehicle, and services expenditures (up to 14 percent for the motor vehicle category). Moreover, the Conference Board's question one is a strong predictor of durable goods (excluding motor vehicles) spending, yielding an incremental adjusted R^2 of 18 percent. In addition, for both indexes, questions three, four, and five hold predictive power for motor vehicle expenditures. The Conference Board's questions three through five also help explain total expenditures.[17]

From the results in Table 3, we arrive at several general conclusions about the types of questions that have significant forecasting ability. First, questions that ask specifically about job prospects in the respondent's area (questions two and four of the Conference Board survey) generally have the most explanatory power.

Second, questions that ask about either the present or the future have more forecasting power than questions that compare the present with the past. The Michigan index's question two, the only question in either index that asks about conditions today relative to the past, has virtually no explanatory power.

Third, questions that ask about consumers' personal financial situations exhibit more predictive power than questions that ask about present buying conditions: for both surveys, the question on personal finances (question five) is significant for some categories of expenditure. The only question about current buying conditions, question one of the Michigan index, elicits virtually no incremental information.

Table 3
PREDICTIVE POWER OF THE SURVEYS' COMPONENT QUESTIONS

Real Personal Consumption Expenditures	Michigan Index		Conference Board Index	
	Present Conditions Component		Present Conditions Component	
	Question 1	Question 2	Question 1	Question 2
Total	-0.01	-0.03	0.04	0.10
	(0.542)	(0.482)	(0.037)	(0.002)
Motor vehicles	0.06	0.02	0.02	0.14
	(0.066)	(0.191)	(0.023)	(0.001)
Goods, excluding motor vehicles	0.09	0.02	0.05	0.07
	(0.128)	(0.262)	(0.337)	(0.200)
Services	-0.01	0.03	0.06	0.06
	(0.322)	(0.149)	(0.035)	(0.058)
Durables goods, excluding motor vehicles	0.03	0.05	0.18	0.13
	(0.406)	(0.086)	(0.002)	(0.004)
	Expectations Component			Expectations Component	
	Question 3	Question 4	Question 5	Question 3	Question 4	Question 5
Total	-0.03	-0.03	0.02	0.09	0.09	0.06
	(0.488)	(0.395)	(0.179)	(0.001)	(0.001)	(0.012)
Motor vehicles	0.04	0.05	0.06	0.10	0.10	0.03
	(0.107)	(0.040)	(0.017)	(0.002)	(0.003)	(0.115)
Goods, excluding motor vehicles	0.01	0.02	0.03	0.0	0.0	0.03
	(0.675)	(0.832)	(0.407)	(0.506)	(0.762)	(0.105)
Services	-0.02	0.01	-0.02	0.01	0.04	0.08
	(0.617)	(0.339)	(0.502)	(0.470)	(0.087)	(0.010)
Durable goods, excluding motor vehicles	0.02	0.06	0.03	0.0	0.06	0.03
	(0.266)	(0.188)	(0.117)	(0.451)	(0.162)	(0.289)

Source: Authors' calculations.

Notes: Figures in parentheses are *p*-values of the joint significance of the lags of the component question. Hypothesis tests were conducted using a heteroskedasticity and serial correlation robust covariance matrix. The sample covers the period from the first quarter of 1968 to the third quarter of 1996.

Fourth, consumer expectations over long-term horizons may be more informative than expectations over short-term horizons for predicting expenditures on large-ticket items such as motor vehicles.[18] Note that questions three, four, and five of the Michigan index are positively correlated with future spending on motor vehicles: the coefficients on the four lags of question three are jointly significant at the 10 percent level, and the coefficients on the four lags of questions four and five, at better than the 5 percent level. These three questions ask about consumers' expectations over a time horizon of one year or more.

Finally, the differences in the types of questions asked by the two surveys may explain our earlier finding that including both measures of consumer attitudes in the forecasting equation better predicts motor vehicle spending than does the use of either index alone. Using both surveys allows the model to capture simultaneously two aspects of consumer sentiment that appear important to motor vehicle spending: consumer expectations over long-term horizons (the Michigan survey) and consumer expectations about job availability (the Conference Board survey).

Out-of-Sample Tests

Our results so far have been obtained by estimating the confidence-augmented equations over the whole sample period. In this section, we test the ability of the equation to forecast out of sample. These tests indicate that the out-of-sample predictive power of the overall Conference Board index was strong in the 1980s, but that it diminished in the early 1990s. Out-of-sample forecast equations augmented with the Michigan index do not generally improve upon the predictive power of the baseline model in any subperiod.

Out-of-Sample Forecasting Performance

To conduct the out-of-sample forecasts, we compare the forecast accuracy of equation 2 over two nonoverlapping evaluation periods across specifications that include either the overall index of each survey or one of the survey's component questions.

The out-of-sample procedure is as follows: as before, the baseline model specifies consumption growth as a function of four lags of the dependent variable, four lags of the growth in real labor income, four lags of the log first difference in the real stock price index as measured by the Standard and Poor's 500 index, and four lags of the first difference of the three-month Treasury bill rate. We then analyze the out-of-sample forecast error of the confidence-augmented models. The model is first estimated using data from the first quarter of 1968 to the fourth quarter of 1981. We then conduct out-of-sample forecasts for two subperiods: the first quarter of 1982 to the fourth quarter of 1989 and the first quarter of 1990 to the third quarter of 1996. We use recursive regressions to reestimate the model, adding one quarter at a time and calculating a series of one-step-ahead forecasts. The forecasts are evaluated by computing the root-mean-squared error from the set of one-step-ahead forecasts.

{Out-of-sample} tests indicate that the . . . predictive power of the overall Conference Board index was strong in the 1980s, but that it diminished in the early 1990s.

Chart 4 provides a visual impression of the relative forecasting power of each model for four different periods. The chart compares the implied consumption levels of the models using the overall indexes with actual levels obtained during those years. As the chart shows, during several episodes in the 1980s and 1990s, the Conference Board-augmented model predicts a level of consumption that was closer to the actual level than that predicted by either the baseline or the Michigan-augmented models.

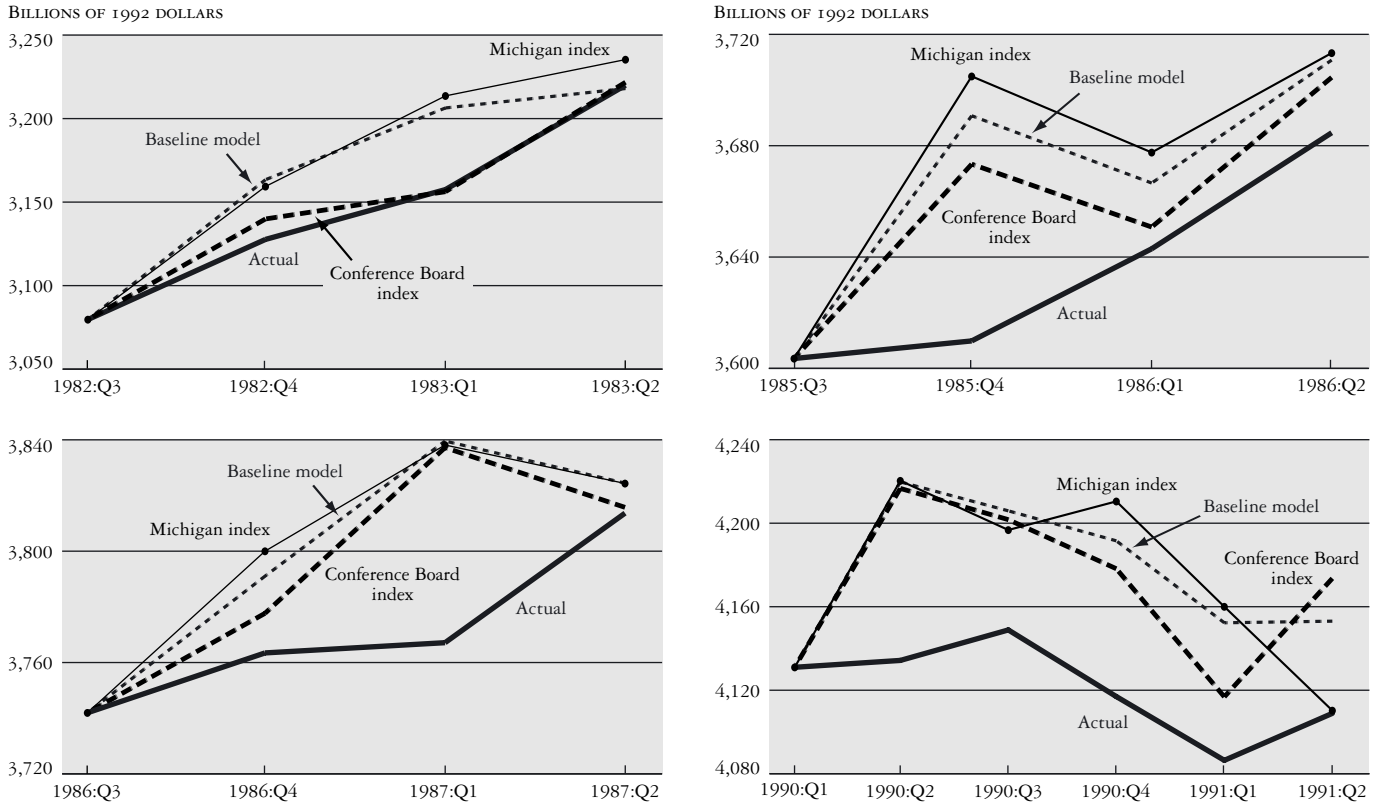
Table 4 summarizes the out-of-sample forecasting performance of each confidence-augmented model. We compare the accuracy of the confidence-augmented equations with that of the baseline model. For each evaluation period and each category of consumer expenditure, the first entry gives the ratio of the root-mean-squared error obtained for the Michigan-augmented model to that obtained for the baseline model.

The second entry gives the ratio of the root-mean-squared error for the Conference Board-augmented model to that obtained for the baseline model. In both cases, results of less than one indicate that using the attitudinal indicator in the forecasting equation improves the out-of-sample forecast relative to the baseline equation. Finally, the third entry gives the ratio of the root-mean-squared error of the Michigan-augmented model to that of the Conference Board-augmented model; a number greater than one indicates that the Conference Board-augmented model outperformed the Michigan-augmented model. The modified Diebold-Mariano test statistic for equal forecast accuracy (see Harvey, Leybourne, and Newbold [1997]) appears in parentheses. We discuss the use of this test statistic below.

For most categories of consumer spending, the forecasting error of the Conference Board-augmented equation is lower than that of the Michigan-augmented equation over most evaluation periods. Moreover, for total personal consumption expenditures and motor vehicle expenditures, the forecasting error of the Conference Board-augmented equation is lower than that of the baseline equation during the 1980s. For example, for total personal consumption expenditures, inclusion of the Conference Board index reduces the root-mean-squared error over the baseline equation by 10 percent for the period from first-quarter 1982 to third-quarter 1996, and by 14 percent for the subperiod from first-quarter 1982 to fourth-quarter 1989. By contrast, the Michigan-augmented equation performs worse than the

Chart 4

Implied Consumption Levels: Actual Relative to Forecast



Source: Authors' calculations.

Note: Dollars are chain-weighted.

baseline model in predicting growth in total personal consumption spending during both of these periods. A similar result holds for the equations that predict growth in motor vehicle spending.[19]

Although the out-of-sample results in Table 4 reveal many qualitative similarities with the in-sample results, a few differences in outcome arise when we estimate the equations over different subperiods. Most notably, while the Michigan index is found to be helpful in forecasting future movements in motor vehicle expenditures when the equation is estimated over the full sample, the out-of-sample results reveal that including the Michigan index improves the predictive power of the baseline equation only in the subperiod from first-quarter 1990 to third-quarter 1996 and weakens the forecasts over the entire first-quarter 1982 to third-quarter 1996 period. Moreover, the out-of-sample predictive power of the Conference Board index appears to be concentrated in the total personal consumption category and in motor vehicle spending. In contrast to the strong in-sample predictive power displayed in Table 2, the Conference Board model does not improve the forecasting performance of the baseline equation in any subperiod for expenditures on goods (excluding motor vehicles).

Table 4
OUT-OF-SAMPLE PREDICTIVE POWER OF ONE-STEP-AHEAD FORECASTS

Real Personal Consumption Expenditures	1982:Q1–1996:Q3	1982:Q1–1989:Q4	1990:Q1–1996:Q3
Total			
Michigan/baseline model	1.014	1.035	1.037
	(0.68)	(0.60)	(0.38)
Conference Board/baseline model	0.900	0.857	1.042
	(-0.70)	(-1.40)	(0.35)
Michigan/Conference Board	1.127	1.208	0.995
	(1.14)	(1.42)	(-0.06)
Motor vehicles			
Michigan/baseline model	1.019	1.029	0.998
	(0.41)	(0.64)	(-0.02)
Conference Board/baseline model	0.930	0.902	0.988
	(-1.17)	(-1.50)	(-0.10)
Michigan/Conference Board	1.096	1.141	1.010
	(1.56)	(1.70)*	(0.11)
Goods, excluding motor vehicles			
Michigan/baseline model	0.990	0.994	1.035
	(-0.25)	(-0.91)	(0.58)
Conference Board/baseline model	1.016	1.013	1.020
	(0.37)	(0.22)	(0.28)
Michigan/Conference Board	0.974	0.981	1.014
	(-0.52)	(-0.80)	(0.25)
Services			
Michigan/baseline model	1.081	1.125	1.030
	(1.87)*	(1.82)*	(0.62)
Conference Board/baseline model	1.029	1.004	1.056
	(0.47)	(0.06)	(0.53)
Michigan/Conference Board	1.051	1.121	0.975
	(0.92)	(1.65)	(-0.30)
Durable goods, excluding motor vehicles			
Michigan/baseline model	1.040	1.024	1.075
	(1.19)	(0.59)	(1.19)
Conference Board/baseline model	1.061	1.088	0.996
	(1.48)	(1.66)	(-0.09)
Michigan/Conference Board	0.980	0.941	1.079
	(-0.41)	(-1.01)	(1.18)

Source: Authors' calculations.

Notes: The table reports the ratio of the root-mean-squared forecasting error. A number less than one indicates that the confidence-augmented model in the numerator has superior forecasting ability. The modified Diebold-Mariano test statistics (Harvey, Leybourne, and Newbold 1997) appear in parentheses. Out-of-sample evaluation periods are reported at the top of each column; the initial estimation period begins with the first quarter of 1968 and ends with the fourth quarter of 1981.

*Significant at the 10 percent level.

The numbers in parentheses in Table 4 give the modified Diebold-Mariano test statistic derived from the method in Harvey, Leybourne, and Newbold (1997) for testing equal forecast accuracy. This statistic has a student's t-distribution and allows the researcher to test whether differences in root-mean-squared error are statistically significant. For each category of consumer expenditure, the statistics indicate whether the out-of-sample forecast error of the confidence-augmented equation is statistically greater than the forecast error of the baseline equation. A positive number indicates that the baseline model has a lower forecast error than the confidence-augmented model. The forecast errors of the confidence-augmented models are also compared with one another; a positive test statistic indicates that the Conference Board-augmented model has a lower forecast error than the Michigan-augmented model.

We report these test statistics but remain skeptical about their value in detecting differences in forecast accuracy. A number of recent papers have documented problems with procedures that test whether differences in out-of-sample forecast error are statistically significant. Researchers often find that variable x Granger-causes variable y in sample, but that out-of-sample tests detect no statistically significant difference in forecast accuracy across the two models

according to whether or not they include x. One possible explanation for differences in in-sample and out-of-sample forecast accuracy is that the in-sample procedure may over-fit the data relative to the out-of-sample procedure. A second possible explanation is that out-of-sample tests simply have little power to reject the null hypothesis of equal forecast accuracy. Clark (1996) shows that tests for equal out-of-sample forecast accuracy generally have much lower power than in-sample Granger causality tests. Thus, the Clark study demonstrates that the discrepancy between in-sample and out-of-sample results may often be attributable to the low power of tests for equal out-of-sample forecast accuracy rather than to true over-fitting in sample. This may explain why we find strong in-sample Granger causality using the Conference Board index and generally no statistically significant difference in the out-of-sample forecasting performance of our models.[20]

Not surprisingly, the test statistics in Table 4 reveal no statistically significant differences in forecast error between the baseline model and the confidence-augmented models for most categories of consumption expenditure over most evaluation periods.

In summary, the results in Table 4 indicate that using the Conference Board index of consumer confidence would have consistently improved out-of-sample forecasts of total or motor vehicle spending growth in the 1980s. After 1990, however, the forecasting power of the model appears to change (Table 4, column 3). In predicting all categories of spending growth except motor vehicles, the baseline model outperforms both confidence-augmented models during this subperiod. Whether the Conference Board index will prove a reliable predictor of consumer spending in the future remains an open question. It is too early to tell whether the forecasting power of consumer confidence displayed by the Conference Board's overall index in the 1980s will return in the late 1990s.

Question-Level Analysis

As a last step, we analyze the out-of-sample forecasting performance of each question over each evaluation period and for every category of expenditure. Because of the large number of results, we present only those combinations for which at least one of the question-level indexes displayed modest improvement in the forecasting power over the baseline model (Table 5).

As Table 5 shows, the best results over the entire period from first-quarter 1982 to third-quarter 1996 are for the confidence-augmented model that uses four lags of the Conference Board's question four on future job avail-

Table 5
Out-of-Sample Predictive Power of One-Step-Ahead Forecasts

	1982:Q1–1996:Q3		1982:Q1–1989:Q4		1990:Q1–1996:Q3	
Real Personal Consumption Expenditures	Michigan	Conference Board	Michigan	Conference Board	Michigan	Conference Board
Total						
Question 1	1.074	0.985	1.117	0.984	1.018	0.987
Question 2	1.002	1.041	1.022	1.069	0.977	1.006
Question 3	0.996	0.955	0.983	0.907	1.011	1.012
Question 4	0.989	0.916	0.980	0.846	1.000	0.995
Question 5	1.006	0.999	1.037	0.913	0.965	1.095
Motor vehicles						
Question 1	0.959	0.957	0.982	0.938	0.908	0.999
Question 2	1.005	0.977	1.039	0.994	0.926	0.940
Question 3	1.016	0.944	1.012	0.918	1.024	1.000
Question 4	0.981	0.930	0.980	0.915	0.983	0.962
Goods, excluding motor vehicles						
Question 1	1.025	0.946	1.033	0.987	1.016	0.900

Source: Authors' calculations.

Notes: The table reports the ratio of the root-mean-squared forecasting error for the equation containing the question to the root-mean-squared forecasting error for the equation without the question; a number less than one indicates that including the question improves the forecast accuracy relative to the baseline model for that particular category of consumption. Out-of-sample evaluation periods appear at the top of each column; the initial estimation period begins with the first quarter of 1968 and ends with the fourth quarter of 1981.

ability. Including this question in the forecasting equation consistently improves the out-of-sample forecasts of total personal consumption expenditure during this period. It also improves the model's out-of-sample performance in both subperiods—most notably in the 1980s. The out-of-sample forecasting power of the Conference Board's question four corroborates the in-sample finding that questions about job availability typically have the most predictive power.

Other results show that the Conference Board's questions one through four generally improve forecasts in every period for motor vehicle expenditure. Michigan's questions one, two, and four are also useful for forecasting motor vehicle spending.

To summarize, like the in-sample tests, the out-of-sample results show that some survey questions have more predictive power than others. Questions that ask about consumers' perceptions of job availability typically have the most explanatory power for future movements in consumption, whereas questions that ask about buying conditions or financial conditions today relative to the past appear to have much less explanatory power.

Conclusion

This paper investigates the impact of consumer attitudes on consumer spending. The purpose of our empirical analysis is to compare the forecasting power of two widely followed measures of consumer perspectives—the Conference Board Consumer Confidence Index and the University of Michigan Index of Consumer Sentiment. We also discuss the ways in which the surveys underlying these measures differ and test whether certain types of survey questions are particularly important for predicting consumer expenditures.

We find that lagged values of the Conference Board Consumer Confidence Index provide information about the future path of spending that is not captured by lagged values of the Michigan Index of Consumer Sentiment, labor income, stock prices, interest rates, or the spending category itself. These results contrast with those of other researchers, such as Carroll, Fuhrer, and Wilcox (1994), who find that consumer attitudes, as measured by the University of Michigan index, contribute little additional information.

The most obvious implication of our empirical results is that forecasts of total personal consumer spending may be made more accurate by utilizing the Conference Board's Consumer Confidence Index. Forecasts are often improved either by replacing the Michigan index with the attitudinal indicator from the Conference Board or by combining the Conference Board data with more conventional economic variables such as income, consumption, and financial indicators.

We also find that the general superiority of the Conference Board index for forecasting consumption appears to be related to the types of questions that make up the survey. The two Conference Board questions that ask specifically about job prospects in the respondent's area exhibit the most predictive power. By contrast, in the Michigan index, the two questions that focus on current buying conditions or financial conditions in the recent past display little incremental forecasting power. Thus, when the surveys of consumer attitudes reveal a major shift in sentiment, policymakers and forecasters might wish to pay close attention to the questions that generated this response. For example, a surge in consumer confidence that is largely driven by the questions about future job availability might suggest greater potential for increased consumer spending than a surge in confidence that is driven by other questions. Consumers seem to spend more when they feel good about future job prospects than they do when they think business conditions are favorable.

We have left at least one important topic for future research: the issue of what theoretical model might account for the spending-confidence correlations we have found. We caution that our results do not prove that consumer attitudes cause changes in consumer spending. Although our analysis explicitly controls for economic fundamentals regarded as important determinants of aggregate consumption growth, the possibility remains that some other variable may be driving the confidence-spending correlations found here. Nevertheless, our results suggest that consumer confidence can help predict consumption, and that consumer attitudes may also act as a catalyst for economic fluctuations.

Appendix A: Correlation Matrix

	University of Michigan Index			Conference Board Index		
	Total	Expectations	Present Conditions	Total	Expectations	Present Conditions
Michigan total	1.00	0.96	0.90	0.69	0.71	0.48
Michigan expectations		1.00	0.75	0.68	0.80	0.42
Michigan present conditions			1.00	0.59	0.45	0.51
Conference Board total				1.00	0.71	0.91
Conference Board expectations					1.00	0.34
Conference Board present conditions						1.00

Appendix B: Definitions of the Estimation Variables

Consumption

We examine five categories of real personal consumption expenditure: total expenditure; motor vehicles; goods, excluding motor vehicles; services; and durables, excluding motor vehicles. The quarterly data are from the U.S. Department of Commerce, Bureau of Economic Analysis.

Labor Income

Labor income is defined as wages and salaries plus transfers minus personal contributions for social insurance. These quarterly components are from the Department of Commerce's National Income and Product Accounts.

Interest Rates

The interest rate is the three-month Treasury bill rate, reported monthly by the Board of Governors of the Federal Reserve System. The data are quarterly averages.

Stock Prices

Stock prices equal the Standard and Poor's 500 composite stock price index (1941-43=10). The data are quarterly averages.

Price Deflator

Nominal labor income and the Standard and Poor's 500 index are deflated by the personal consumption expenditure implicit price deflator (1992=100). The data are reported quarterly in the National Income and Product Accounts. The data reflect revisions in September 1993.

ENDNOTES

The authors are grateful to Jeffrey Fuhrer, Jonathan McCarthy, Patricia Mosser, Gabriel Perez Quiros, Robert Rich, Rae Rosen, Christopher Sims, Charles Steindel, and Egon Zakrajšek for helpful comments, and to Beethika Khan for excellent research assistance.

1. Of course, there may have been other proximate causes of the 1990-91 recession such as the Persian Gulf War and commodity-price or bank-loan supply shocks associated with the war.

2. The Michigan index begins with quarterly data in 1952; the Conference Board index with bimonthly data in 1967.

3. Early investigators of the explanatory power of consumer confidence include Fair (1971), who links the University of Michigan index with both durable and nondurable consumer expenditures, and Mishkin (1978), who argues that the Michigan index may be a good proxy for the consumer's subjective assessment of the probability of future financial distress. More recent work analyzing the Michigan index can be found in Carroll and Dunn (1997), Carroll, Fuhrer, and Wilcox (1994), Fuhrer (1993), Leeper (1992), and Matsusaka and Sbordone (1995).

4. We leave for future research the question whether some theoretical model might explain the predictive power of consumer attitudes for consumption.

5. Because the Conference Board index includes a question about nominal income, it may overstate "confidence" during periods of high inflation.

6. This difference in time horizons may have some effect on response patterns and hence on index results.

7. There may be some sample selection bias in both surveys, but any such bias is assumed to be constant over time and so has virtually no effect on the indexes.

8. Because of differences in index construction, discussed earlier, the Conference Board's index has a wider range of movement than Michigan's. However, on a standardized basis, the Conference Board's index is significantly *less* volatile—that is, it has a higher signal-to-noise ratio than Michigan's index.

9. As noted earlier, the University of Michigan quarterly data are available from 1952, while the Conference Board data do not begin until the first quarter of 1967. To maintain a basis of comparison across regressions, we use the largest possible sample for which both indexes are available.

10. Estrella and Hardouvelis (1989) have established the forecasting power of this "term structure" spread for several real variables.

11. The growth in spending on durable goods may be positively autocorrelated, with the error term following a first-order moving-average process (see Mankiw [1982]). First-order autocorrelation in the error term may cause the error term to be correlated with the one-period-lagged endogenous variable, a condition that could skew in-sample statistical tests of the joint marginal significance of the explanatory variables (the reported *p*-values). To address this problem, we explicitly model the error term, ε_t, following an MA(1) process in the in-sample regressions. This strategy is derived from Carroll, Fuhrer, and Wilcox (1994). Allowing for an MA(1) in the error term requires nonlinear estimation, and we use nonlinear least squares in our in-sample estimation of equation 1 and in the confidence-augmented equations that follow. The coefficient on the lagged-moving-average term generally has the expected negative sign. For example, for total real personal consumption expenditures and the confidence-augmented equation using the Michigan index, the coefficient is estimated at -0.8, with a standard error of 0.1. See Carroll, Fuhrer, and Wilcox (1994).

12. We do not report results for the present conditions component because preliminary tests indicated that the expectations component of both indexes typically exhibited greater forecasting power.

13. Previous research (for example, Leeper [1992]) suggests that consumer confidence may be linked to economic indicators such as unemployment and industrial production largely because of unusually volatile movements in consumer attitudes during the Persian Gulf War and the 1990-91 recession. To control for this possibility, we include a dummy variable set equal to one in the quarters corresponding to the 1990-91 recession. We then eliminate the dummy variable and perform out-of-sample forecasts over several evaluation periods using the beginning of the 1990-91 recession as a break.

14. Adding a dummy variable for the third quarter of 1980 to account for credit controls does not significantly alter the results; in the sample controlling for the 1990-91 recession, the incremental adjusted R^2 is .09, and the lags of the Conference Board index are jointly significant at better than the 1 percent level. The adjusted R^2 from a regression of total personal consumption expenditure growth on the controls alone is approximately .40.

15. By regressing consumption growth on four lags of the overall index, we implicitly restrict the coefficient on each component (relative to its share in the overall index) to be the same. One question to consider is whether the forecasting power of the Conference Board index would be improved by regressing the consumption growth category on the

Note 15 continued
expectations and present conditions components separately. We investigated this question but found that the incremental adjusted R^2 increased significantly in just one category: motor vehicle spending. In that equation, when we added four lags of each component separately, the increment to the adjusted R^2 increased to 16 percent, from 5 percent.

16. Unlike the overall indexes, however, the question-level indexes are not pegged to a base year. Question-level data for the University of Michigan survey come from the Board of Governors of the Federal Reserve System. We thank Lynn Franco of the Conference Board for providing us with data on the Conference Board questions.

17. Conference Board question one also has statistically significant explanatory power in the motor vehicle expenditure equation. However, the increment to the adjusted R^2 is quite modest and considerably smaller than that produced by the other questions for this expenditure category. Michigan's question one is a statistically significant predictor of motor vehicle spending at the 10 percent level but not at the 5 percent level.

18. This finding makes sense because motor vehicles are more likely to be financed using long-term credit than are other durable goods.

19. Note that the first subperiod does not include the 1990-91 recession, so that the recession cannot explain the predictive power of the Conference Board index.

20. There are other problems with statistical tests for equal forecast accuracy. Harvey, Leybourne, and Newbold (1997) have documented the severe size problems of the standard Diebold-Mariano test. Their modified test, used in this study, goes part of the way toward fixing the size problems but does not eliminate them.

References

Bram, Jason, and Sydney Ludvigson. 1997. "Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race." Federal Reserve Bank of New York Research Paper no. 9708.

Campbell, John Y., and Gregory Mankiw. 1989. "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence." In O. Blanchard and S. Fischer, eds., NBER MACROECONOMICS ANNUAL. Cambridge: MIT Press.

Carroll, Christopher D., and Wendy Dunn. 1997. "Unemployment Expectations, Jumping (S,s) Triggers, and Household Balance Sheets." NBER MACROECONOMICS ANNUAL. Cambridge: MIT Press.

Carroll, Christopher D., Jeffrey C. Fuhrer, and David W. Wilcox. 1994. "Does Consumer Sentiment Forecast Household Spending? If So, Why?" AMERICAN ECONOMIC REVIEW 84, no. 5: 1397-1408.

Clark, Todd E. 1996. "Finite-Sample Properties of Tests for Forecast Equivalence." Federal Reserve Bank of Kansas City Research Paper no. 96-03.

Estrella, Arturo, and Gikas A. Hardouvelis. 1989. "The Term Structure as a Predictor of Real Economic Activity." Federal Reserve Bank of New York Research Paper no. 8907.

Fair, R. C. 1971. "Consumer Sentiment, the Stock Market and Consumption Functions." Princeton University Econometric Research Program Research Memorandum 119.

Flavin, Marjorie. 1981. "The Adjustment of Consumption to Changing Expectations about Future Income." JOURNAL OF POLITICAL ECONOMY 89, no. 5: 974-1009.

Fuhrer, Jeffrey C. 1993. "What Role Does Consumer Sentiment Play in the U.S. Economy?" Federal Reserve Bank of Boston NEW ENGLAND ECONOMIC REVIEW, January-February: 32-44.

Harvey, David, Stephen Leybourne, and Paul Newbold. 1997. "Testing the Equality of Prediction Mean Squared Errors." INTERNATIONAL JOURNAL OF FORECASTING 13, no. 2 (June): 281-91.

Leeper, Eric M. 1992. "Consumer Attitudes: King for a Day." Federal Reserve Bank of Atlanta ECONOMIC REVIEW 77, no. 3 (July-August): 1-16.

Mankiw, Gregory. 1982. "Hall's Consumption Hypothesis and Durable Goods." JOURNAL OF MONETARY ECONOMICS 10, no. 3: 417-25.

Matsusaka, John G., and Argia M. Sbordone. 1995. "Consumer Confidence and Economic Fluctuations." ECONOMIC INQUIRY 33, no. 2 (April): 296-318.

Mishkin, Frederic S. 1978. "Consumer Sentiment and Spending on Durable Goods." BROOKINGS PAPERS ON ECONOMIC ACTIVITY, no. 1: 217-32.

Wilcox, David W. 1992. "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work." AMERICAN ECONOMIC REVIEW 82, no. 4: 922-41.