DOES CONSUMER CONFIDENCE FORECAST HOUSEHOLD EXPENDITURE? A SENTIMENT INDEX HORSE RACE

Jason Bram and Sydney Ludvigson

Federal Reserve Bank of New York Research Paper No. 9708

March 1997

This paper is being circulated for purposes of discussion and comment only.

The contents should be regarded as preliminary and not for citation or quotation without permission of the author. The views expressed are those of the author and do not necessarily reflect those of the Federal Reserve Bank of New York of the Federal Reserve System.

Single copies are available on request to:

Public Information Department Federal Reserve Bank of New York New York, NY 10045

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

By Jason Bram and Sydney Ludvigson*

Draft: March 7, 1997

Abstract

This paper investigates the forecasting power of consumer sentiment for household expenditure. We compare the predictive power of two measures of consumer attitudes, and then further compare each survey's expectations component with one another, and with the broader sentiment measures. The results indicate that lagged values of the Conference Board's overall confidence and expectations measures have stronger incremental predictive power for more categories of the growth in consumption expenditure than do measures available from the University of Michigan, though the latter is correlated with future expenditures on automobiles even after economic fundamentals are controlled for. The paper also discusses structural differences between the two surveys and suggests how these differences might be related to the discrepancy in forecasting power.

*Domestic Research Function, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045. We are grateful to Jonathan McCarthy, Patricia Mosser, Gabriel Perez, Rob Rich, Rae Rosen, Charles Steindel, and Egon Zakrajsek for helpful comments, and to Beethika Khan for excellent research assistance. The views expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the Authors.

1. Introduction

The behavior of consumer attitudes is a subject of enduring captivation to policy makers and economic forecasters. Household sentiment has been cited as the leading cause of the 1990-1991 recession¹, and historically high levels of current confidence measures have helped to fuel speculation that the economy may be heading for a period of overheating (Figure 1).

Do consumer attitudes help forecast the future of consumer spending? Can household mood swings tell us anything about real activity that is not already captured by economic fundamentals? If so, which measure of these attitudes contains more information?

This article addresses the questions raised in the preceding paragraph by investigating the role consumer sentiment plays in predicting various categories of household expenditure. The literature to date has focused almost exclusively on whether consumer attitudes are related to household expenditure when the former are measured by the University of Michigan's overall Consumer Confidence index.² In contrast, this study compares the forecasting power of two separate measures of consumer attitudes, by using surveys of consumer confidence available from both the University of Michigan and the Conference Board. We further break each of those indexes down into their respective consumer expectations components, and separately assess the forecasting power of those attitudinal subcategories.

We find the forecasting power of consumer sentiment depends on which measure is used. Measures available from the Conference Board (CB) indicate that consumer perspectives have both economically and statistically significant explanatory power for four of five spending categories we consider, even when other economic indicators are controlled for. In contrast, measures available from the University of Michigan's Survey Research Center (MI) exhibited a much smaller forecasting role for most categories of consumer spending, though its overall index does appear to

¹For example, this suggestion can be found in Carroll, Fuhrer, and Wilcox (1994).

² Early studies on the explanatory power of consumer sentiment include Fair (1971) who found an association of the Michigan index with both durable and nondurable consumer expenditure, and Mishkin (1978) who argues that the index of consumer sentiment may be a good proxy for the consumer's subjective probability of future financial distress. More recent work includes studies by Carroll, Fuhrer, and Wilcox (1994), Fuhrer (1993), Leeper (1992), and Matsusaka and Sbordone (1995).

be related to future changes in automobile consumption.

A relevant question for the economic forecaster is whether sentiment has any predictive power for future changes in consumption apart from the information contained in other economic variables. Using simple reduced-form forecasting equations, we conclude that the CB measure has significant *incremental* predictive power for several categories of consumer spending, and that both the CB and MI measures have incremental predictive power for automobile expenditures. In other words, measures of consumer attitudes appear to contain information about consumer spending that is not captured by "economic fundamentals" like income, interest rates, and stock prices.³

Existing research which studies the relation of consumer sentiment to real variables other than consumption expenditure suggests that correlations between sentiment and real activity may be significantly compromised by eliminating the unusually volatile movements in confidence associated with the Gulf War recession. For example, Leeper (1992) finds such evidence in testing whether consumer attitudes influence unemployment and industrial production. Our results also indicate that the forecasting power of consumer attitudes is somewhat sensitive to whether the 1990-'91 recession is omitted from the statistical analysis, though results using the MI index were more generally dependent on this omission than were results using the CB index. In short, our findings suggest that, at least for some categories of consumer expenditure, the incremental information contained in the CB index is not simply the result of an isolated episode associated with the Gulf War.

The rest of this paper is organized as follows. To provide background information about the sentiment measures we consider, the next section discusses in detail the structural and institutional differences between the Michigan and Conference Board surveys. The goal of this section is to provide a frame-of-reference for thinking about how differences in survey methodology across attitudinal measures might contribute to the dissimilar statistical relationships found in the data. The primary focus of our empirical analysis is to contrast the incremental information content of the two

³We remain agnostic as to what theoretical model this predictive power might be consistent with. One possibility is that sentiment simply forecasts the true state of the economy and plays no causal role in determining real spending. Alternatively, it could be a true underlying causal factor for spending, a possibility recently raised by Matsusaka and Sbordone (1995) who argue that sentiment may play such a role in models which exhibit multiple Nash equilibria with strategic complementarities.

sentiment measures; we take this up in section 3.1 with a discussion of simple reduced form regression results. Finally, for comparison with previous studies on consumer sentiment and its relationship to real variables, section 3.2 briefly addresses a separate issue, that of summarizing the historical dynamics between sentiment and consumer spending. Section 4 is a summary and contains concluding remarks.

2. Differences between the Michigan and Conference Board Indexes

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 (Figure 1). Michigan's (MI) survey was developed shortly after WW2 (late 1940s) and, at first, was conducted annually. It was expanded to a quarterly survey in 1952, and to a monthly survey in 1978. The Conference Board's (CB) survey, which was modeled on Michigan's (though with some key differences, discussed below), was developed in 1967 as a bimonthly survey; it was expanded to a monthly series in June 1977. All the data used in this analysis are quarterly averages.

While academic studies have focused almost exclusively on the MI index, the financial markets follow both indices closely⁴. The media and the general business community tend to focus more on CB's, in part because it is issued in an official release and is reported in greater detail. While the two measures generally follow similar trends over the long run, they sometimes give opposite signals on a month-to-month basis. For this reason, having both measures available is helpful to those forecasting the short-term economic outlook. Although they broadly measure the same concept, there are some fundamental differences between the two indices, and understanding these differences is essential in interpreting the two series and identifying signals. The major difference between the two surveys involves the actual surveys questions. In addition there are technical differences regarding survey methodology and index computation methodology.

2.1 Survey Questions—Present Conditions vs. Expectations

⁴Virtually all published academic research on this topic focuses on MI's index (mostly due to its longer history). However, financial news-wire services tend to focus on CB's index for pre-release commentary and forecasts (due to its more official and detailed releases).

Both indices are based on responses to five survey questions. In both surveys, two of the five questions gauge respondents' assessment of **present conditions**; the responses are used to compile a *Present*⁵ Index, which gets a 40% weight in the overall index. The other three questions gauge consumers' **expectations**; these responses are used to compile an *Expectations* index, which has a 60% weight.

The *Present Conditions* questions are fundamentally different in the two surveys: The questions differ somewhat in focus—CB's survey focuses on business conditions and job availability, while MI's asks about big-ticket household purchases and changes in the respondent's financial situation (Figure 2). As a result, CB's survey is more closely associated with labor market conditions, while MI's tends to be influenced more by interest rates and inflation.

In addition, the surveys differ over how consumers measure their current situation. CB's survey uses an absolute benchmark for comparison on both questions, while MI's uses a relative benchmark ('a year ago') on Q2. In other words, MI's present conditions index blends consumers' perception of *level* (Q1) and *change* (Q2) in economic activity, while CB's just reflects their perception of *level*. These differences are reflected in the cyclical behavior of the respective indices: MI's *Present Conditions* Index tends to peak in the early stages of economic recovery, when growth peaks and interest rates and inflation are low, while CB's *Present Situation* Index generally peaks in the late stages, when unemployment is lowest and the level of economic activity is highest (Figure 3).

The *Expectations* questions are somewhat more comparable in the two surveys: CB's survey asks about expected business conditions and respondent's expected income ⁶; MI's poses parallel questions—on expected business conditions and respondent's expected financial situation. However, MI also includes a query on the nation's long-term economic prospects, while CB substitutes a question about expected job availability (Figure 4).

As was the case in the current situation questions, the surveys differ over how consumers

⁵We use the term *Present* as a generic reference to the two Indices—MI's *Current Conditions* Index and CB's *Present Situation* Index.

⁶ Since CB includes a question about (nominal) income, that index may overstate "confidence" during periods of high inflation.

measure their expectations. First, the two surveys specify different time horizons for expectations: CB uses 6 months on all three queries, while MI uses one year on two of its questions and five years on the third. However, there is little evidence that consumer expectations are much affected by the specific time horizon specified. Moreover, while CB's questions all focus on expected *change* (i.e. "better", "worse") MI's blend the concepts of *change* and *level* (i.e. "good", "bad").

Despite these differences, the two *Expectations* indices are much more highly correlated with each other (82%) than the two *Present* indices (49%) (Figures 3, 5). Moreover, the *Present* and *Expectations* component indices are more closely correlated with one another in MI's survey (73%) than in CB's (40%). In other words, CB's Present index tends to follow a substantially different path than the other three component measures. Because CB's Present Situation Index reflects the level of economic activity⁷, while CB's Expectations Index reflects change, it may be best to observe them as two separate measures. This is not as clear-cut for MI's components, since each one reflects perceptions of both change and level of economic activity. Another difference between the two surveys is that CB's indices tend to track labor markets, while MI's tend to be influenced more by inflation and credit conditions (interest rates)— this is particularly true of the *Present* indices.

2.2 Survey Methodology—How and when the surveys are conducted

Aside from differences in the survey questions, there are also some differences in surveying techniques and timing. MI conducts its survey by phone, with calls made throughout most of the month. MI's sample size is 500, though the preliminary (mid-month) release is based on about 250 interviews made during the first 8-10 days of the month. CB sends out a mail survey at the end of the prior month; responses flow in throughout the month (mostly during the first 2 weeks, then tapering off). TCB's sample size is roughly 3,500 (of a total mailing of 5,0008), with preliminary figures based on about 2,500 responses received over the first 16-18 days of the month.

⁷ For example, the level of TCB's Present Situation Index is extremely correlated with the level of the unemployment rate, and 12 month change in the index is a leading indicator of 12-month (payroll) employment growth.

⁸ 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 indices.

These differences in survey timing affect the release schedules. Both organizations issue preliminary reports based on early returns and later issue "revised" or "final" figures, based on the full month's responses. MI makes its preliminary indices available (though not through an official release) on the second or third Friday of the month, based on phone interviews conducted in the first 1-2 weeks of the month; final figures are generally made available on the last Friday of the month. CB issues a formal release of its preliminary figures on the last Tuesday of the month, based on returns received during the first 2-3 weeks of the month; revised (final) data are released a month later—along with the next month's preliminary figures. There are no subsequent revisions to either series, except when CB revises its historical seasonal factors every few years. In summary, MI's indices are available about two weeks earlier than CB's but are based on a much smaller sample.

2.3 Index Methodology—How the raw responses are transformed into a single index

Yet more differences between the MI and CB measures involve the specific statistical methodologies and formulas used to compile the indices from the raw response data.

The first step involves distilling the three response percentages into a single diffusion measure for each question. Both surveys offer positive, negative and neutral response options on all five questions; MI calculates its diffusion measure based on an arithmetic formula [100+p-n], while CB uses a ratio formula [p/(p+n)]. In essence, CB's methodology implicitly allocates the neutral reposes in proportion to the positive and negative responses; in contrast, MI's methodology implicitly allocates neutral responses evenly between positive and negative.

The next step involves averaging together the question-specific diffusion measures and converting them into base-year indices (by dividing the current value by the average value for the benchmark year—1966 for MI; 1985 for CB). However, one methodological difference is that CB converts the diffusion measures to base-year indices *before* averaging them together, while MI makes the base-year conversion *after* averaging the diffusion measures⁹.

⁹ This represents a statistical quirk in CB's methodology: because the diffusion measures are converted into a base-year indices *before* they are averaged arithmetically, a given question's effective weight in the index is influenced by the selection of the base year. Hypothetically, a month-to-month rise in the overall index can become a decline simply as a result of changing the base year. This is similar to problems associated with the old fixed-base-year GDP deflator. In

Finally, CB's index is seasonally adjusted ¹⁰ while MI's is not; however, there is not much seasonality in either measure, and CB's seasonal adjustment does not materially change that series. The main result of these methodological differences is that CB's index has a wider range of movement than Michigan's—a good rule of thumb is that a 1 point move in Michigan's index is comparable to a 2-point move in CB's.

2.4 Interpreting the Indices

While both indices measure consumer confidence, it is important to understand the respective methodologies in order to interpret them properly. The most basic issue is how to interpret level versus change. On CB's index, it is useful to examine the Present and Expectations components separately—the former as a proxy for the level of economic activity, the latter as a proxy for growth; on MI's, this distinction is less clear, as both components mix level and growth measures. One must also be careful in comparing moves in the MI and CB indices; due to differences in methodology and base years, monthly changes should not be compared in absolute terms but rather on a standardized basis (adjusting for range of movement)—a 1-point change in MI's index is roughly comparable to a 2-point change in CB's. There is no basis at all for comparing index levels.

One advantage of MI's index is its earlier availability; however, one should bear in mind that MI's preliminary figures are based on a much smaller sample size (about 250) than CB's (about 2,500), and are thus subject to somewhat greater measurement error. Finally, because the two questionnaires focus on different issues, the indices sometimes give mixed signals; during such periods, having both indices is particularly useful, because their contrasting signals may offer some insight on how various segments of the economy are performing.

practice, however, this quirk has no discernible effect on the quality of CB's index.

¹⁰ Another quirk in CB's index: raw response percentages are seasonally adjusted using multiplicative (rather than additive) seasonal factors; then these adjusted percentages are used to compile the overall indices (rather than computing indices first and then seasonally adjusting). In practice, though, the seasonal adjustment process has little effect, since the officially reported (s.a.) Index series is not significantly different from a parallel index based on unadjusted data.

3. Empirical Results

This section presents the results of our empirical investigation to determine how well different sentiment measures predict consumer spending. Section 3.1 contains the main focus of our analysis, and presents the outcome of simple reduced-form forecasting regressions for consumption growth; section 3.2 briefly discusses the findings from a vector autoregression analysis. In both sections, we consider the effect of sentiment on five separate 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.¹¹ An appendix reviews the data.

3.1 Reduced-Form Regressions

One way to assess the predictive ability of consumer sentiment for spending is to examine reduced-form evidence from simple forecasting equations for consumption growth. To do so we consider regressions of the form,

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

where, C_i stands for real consumption spending, S is the measure of consumer sentiment or consumer expectations, and Z_i is a vector of control variables. Nonlinear least squares (NLLS) is

¹¹The disaggregation of total expenditure into the first four of these categories has been advocated by Wilcox (1992), who argues that those categories better reflect the procedures used by the Bureau of Economic Analysis to estimate consumer spending than do other categorizations of consumption, such as, for example, the more conventional disaggregation of consumption into nondurables and services and durables categories.

¹²The error term, ϵ_t , is assumed to be distributed MA(1). Because most of the spending categories we consider contain durable goods, the growth in spending is *a priori* expected to be positively autocorrelated, and the error term should follow a first-order moving average process (see Mankiw, 1982). Ludvigson (1996) documents the importance, for obtaining consistent parameter estimates, of explicitly modeling the error term as a moving average process in this form of forecasting equation. The appropriate estimation technique in this case is Nonlinear least squares which we use below.

used to estimate the forecasting equation in quarterly data with the number of lags, *n*, equal to 4. The sample period is 1959:1 to 1996:3 when using the Michigan index, and 1968:1 to 1996:3 when using the Conference Board index.¹³

Table 1, top panel, shows the results of estimating equation (1) without control variables when S is measured by the Michigan overall sentiment index (left columns), and the Conference Board's overall confidence index (right columns). The bottom panel shows analogous figures for the Michigan and Conference Boards' expectation components. The numbers in each cell are the adjusted R^2 statistic from a regression of the growth in the spending category on lags 1 through 4 of the sentiment measure taken alone; numbers in parenthesis are the p-values for the joint marginal significance of the four sentiment lags. The p- values give an indication of how statistically important the sentiment variables are. A number equal to 0.01, for example, says that there is a probability of 1 percent or less that the explanatory power was generated merely by chance.

As Table 1 shows, lagged values of the consumer sentiment index taken on their own explain a substantial portion of the one quarter ahead variation in consumer spending for every category of spending considered. This is true for every measurement index considered. For example, both the CB index and the MI index of overall sentiment explain about 15 percent of the variation in total, personal consumer spending one quarter hence. The CB expectation index explains an even larger fraction of the regression variance for total PCE spending, equal to about 24 percent. These findings are strongly statistically significant whether or not the 1990-'91 recession is eliminated from the sample.

While the findings in the first two columns of each table show that sentiment, by itself, helps to predict the future course of consumption, a more important question is whether sentiment contains any information not captured by economic fundamentals. Are the sentiment lags still significant when the control variables are included in (1)? And if so, is the explanatory power added of notable magnitude?

¹³The discrepancy in reduced form forecasting power which we report below does not appear to be an artifact of the different sample periods over which the regressions are performed; similar results obtained from running the regression with the Michigan Index over the smaller sample period for which the Conference Board index is available.

Table 2 shows statistical results from running the regression on both the sentiment lags and a set of control variables. Adding the control variables allows us to ask whether sentiment has any *incremental* explanatory power once these economic fundamentals are accounted for. The control variables included in Z_i are four lags of the dependent variable, four lags of the growth in real labor income, four lags of log first difference in the real stock price index, and four lags of the first differenced three month t-bill rate. One way to asses the information content of consumer sentiment is to record the *increment* to the adjusted R^2 from adding the sentiment lags to the regression of consumer spending on the control variables, Z. These values are given in Table 2 for the MI and CB overall indexes (top panel) and the MI and CB expectations components (bottom panel). The numbers give the increase in the fraction of one-quarter ahead variation in consumption growth explained from adding the sentiment lags to the equation. Numbers in parenthesis are the p-values for the joint marginal significance level of the sentiment lags.

Table 2 shows that, once the 1990-'91 recession is eliminated from the sample and economic fundamentals are controlled for, lagged values of the MI measure of consumer sentiment add nothing to the adjusted R^2 in a regression with total, personal consumption growth as the dependent variable; the impact is both economically and statistically not different from zero. The incremental adjusted R^2 from the regression using the MI expectations component is substantially higher, but the p-value indicates it is not a statistically significant increment.

In contrast, both the CB measure of consumer confidence, and its measure of consumer expectations, appear to be incrementally informative about future consumer spending growth. Even

¹⁴ Carroll, Fuhrer, and Wilcox (1994), estimated the type of forecasting equation given in (1) using lagged values of the dependent variable and of labor income growth as controls. However, one reason sentiment may be meaningful is because it is made available on a more timely basis than other economic indicators. Financial market indicators are available on an almost continuous basis, and some authors (e.g. Leeper, 1992) have argued that sentiment is only weakly correlated with variables like unemployment and industrial production once financial indicators are controlled for. To assess this possibility for consumer spending, we include the log first difference of the real stock price and the first difference of the 3 month T-bill rate as additional controls in (1). We also tried using the unemployment rate in place of labor income growth, and the spread between the ten year T-bond rate and the one year T-bill rate in place of the first difference of the three month T-bill rate. These changes did not qualitatively alter the conclusions.

controlling for the 1990-'91 recession, adding four lags of the CB confidence index to the regression equation (1), increases the fraction of regression variance explained by 9 percent using the over all index, and by 12 percent using the expectations component. Both outcomes convey a very high degree of statistical confidence.¹⁵

Results in the second row in Table 2 suggest that both overall indexes display some incremental predictive power for automobile spending. Lagged values of the MI sentiment index increase the fraction of the growth in motor vehicle spending explained by 2 percent after the control variables and 1990-'91 recession are accounted for, a small amount but marginally statistically significant. Adding lags of the MI expectation component also increases the fraction of regression variance explained by 2%, but this value is not statistically significant. On the other hand, the CB measures have a much stronger incremental impact on the growth in automobile spending; for example, the overall index increases the adjusted R^2 by 5 percent while the expectations component increases it by 10 percent. Both increments to the forecasting power are strongly statistically significant using the Conference Board indexes.

For other categories of spending, lagged values of the either the MI index of consumer sentiment or its expectations component, in general, add relatively little explanatory power to the consumption growth regressions. ¹⁶ For example, for services spending growth, the increment to the fraction of regression variance explained is often quite small and the result is not statistically significant. In addition, neither MI measure is related to future changes in durables excluding motor vehicles. An important exception is goods excluding motor vehicles: interestingly, the MI

¹⁵Adding a dummy variable for the quarter 1980:3 (credit controls) did not significantly alter the results; in the sample controlling for the 1990-'91 recession, the incremental, adjusted R^2 is 0.9, and the lags of the CB 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 about 0.40.

¹⁶The statistical significance of the Michigan *Expectations* component for total PCE expenditures seems to be somewhat sensitive to how the total PCE is deflated; we experimented with using the chain weighted deflator as well as the implicit price deflator (the results of which are reported above) and found a stronger statistical relationship between the expectation index lags and total consumption growth. However, the incremental adjusted R^2 using either index remained quite small.

expectations component is found to be strongly related to future changes in this spending component once the 1990-'91 recession is controlled for.

Lagged values of the CB overall confidence index appear to be of value in predicting several other categories of consumer expenditure. The increment to the adjusted R^2 for services using the CB index is about 2 percent, a small figure but marginally statistically significant. However, for durables excluding motor vehicles, adding lags of the overall CB index increases the fraction of regression variance explained by a large 15 percent, a finding that is strongly statistically significant. In contrast, the CB expectations component appears to be strongly related to future services expenditure (increasing the adjusted R^2 by a statistically significant 6%), but only weakly related to future expenditures on durables excluding motor vehicles (causing a negative value in the incremental adjusted R^2). However, neither CB measure is statistically related to the future growth in spending on goods excluding motor vehicles.

A comparison across columns indicates the predictive power of both confidence measures is somewhat sensitive to whether the 1990-'91 recession is controlled for, though in general this appears to be more true for the MI measures. For example, controlling for this portion of the sample period eliminates the forecasting power of lagged values of both MI measures for the growth in total PCE spending. It also substantially weakens the association of the overall MI index with the growth in future automobile spending, decreasing the increment to the adjusted R^2 from 11 to 2 percent. Interestingly, there is one category of consumer spending for which controlling for the 1990-'91 recession has the opposite effect: the MI expectations component has significant incremental forecasting power for the future growth in goods excluding motor vehicles only after the 1990-91 recession is accounted for. In general, the forecasting power of the CB measures appear to be much less sensitive to whether this portion of the sample is controlled for.

In summary, results from estimating the reduced form regressions in (1) illustrate several important points. First, both the CB index of consumer confidence, and its index of consumer expectations, help forecast future growth in consumer spending for several expenditure categories. The overall CB measure appears to be related to four of five categories we consider, and its expectations component is related to at least three of the five. Second, this forecasting power of the Conference Board measures are not only statistically significant, but also economically meaningful:

adding lagged values of the CB indexes to the regression increases the fraction of the growth in spending explained by 7 to 15 percent over that which arises from regressing spending growth on the control variables alone. Third, unlike the MI index, the CB confidence measures are not, in general, highly sensitive to the exclusion of the 1990-'91 recession from the sample. Fourth, these results contrast with those from regressions using the MI measures. Of the five spending categories we consider, the overall MI index appears to have incremental explanatory power only for the growth in motor vehicle spending, while the MI expectations component has incremental explanatory power for the single category of goods excluding motor vehicles. In addition, the overall magnitudes of the incremental explanatory power are smaller than those found when using the CB measure, with the highest incremental adjusted R^2 equal to 8 percent.

The discussion in Section 2 may hint at why the sentiment surveys appear to differ so strikingly in their incremental information content. Though a formal investigation to determine how various distinguishing characteristics of the survey methodologies may have contributed to the empirical outcome is beyond the scope of this paper, discrepancies in the survey questions identified in Section 2 may at least be suggestive. For example, the Michigan present conditions questions appear to be a mix of "levels" questions (is now a good time to buy major household items?) and "growth" questions (are you better off financially today compared to a year ago?), while the MI expectations questions appear to be primarily level questions about the future. The two components taken together imply that overall, the MI questionnaire focuses more heavily on the *level* of economic activity, both today and in the future. In contrast, the Conference Board's present condition questions center on the level of economic activity (how good are conditions today?), while its expectations questions emphasizes the growth in economic activity (how are conditions today relative to six months from now?). This suggests the overall CB index focuses more heavily on the *growth* in economic activity, since the expectations component receives 60 percent of the weight.

Divergence in the surveys' questionnaires may be relevant because the reduced form results presented here are for the *growth* in consumer spending. Note that the CB expectation component is, in general, more highly correlated with future changes in spending than is the overall index. This suggests the expectations component, rather than the levels-focused present conditions component, may comprise the bulk of the predictive power in the overall CB index. If consumers' perceptions

about the growth in economic activity (as opposed to their perceptions about the current or future *level* of economic activity) are more important for determining the growth in their spending, this could explain why lagged values of the CB expectations component have more predictive power for spending growth than do lagged values of the overall index, and why lagged values of either MI measure (both of which seem to concentrate on questions about levels) appear to be more weakly related to current spending growth than is either CB measure.¹⁷

3.2 Vector Autoregression Analysis

The last section analyzed the predictive power of consumer attitudes in reduced form regressions. The reduced form regressions are useful for evaluating whether consumer attitudes have any informational content above and beyond that found in other economic indicators. This section briefly turns to a different topic, that of summarizing the dynamic relationship between consumer spending and consumer sentiment.¹⁸

To analyze the dynamic impact of sentiment on spending it is useful to employ an alternate estimation strategy. In this section, we estimate a vector autoregression (VAR) consisting of one equation for each variable that was included in the reduced form analysis above.¹⁹ It is important

¹⁷ This hypothesis is not so obviously applicable to the question of why the MI index appears to forecast automobile expenditures and not other spending categories, or why its explanatory power appears to be sensitive to the inclusion of a 1990-1991 recession dummy. However, it's conceivable that households may have perceived the Gulf War as a rare transitory shock to the level of economic activity, in which case, during this period, individuals may have been more apt to approach questions about their economic security in terms of a comparison with more ordinary times, rather than as a query about the level of conditions today or in the future.

¹⁸ Several existing papers have made use of vector autoregressions to analyze the dynamic impact of the Michigan Index of Consumer Sentiment on real variables; see for example, Leeper (1992), Fuhrer (1993), and Matsusaka and Sbordone (1995).

¹⁹We estimate three and five variable VARs. The data are quarterly; likelihood ratio tests indicated that the appropriate lag length was five. Quantity variables are in logs. They are placed in the following order in the three variable VAR: sentiment, the log of real consumer spending, and the log of real labor income. Leeper (1992) and Fuhrer (1993) have emphasized the importance of including financial variables in the VAR. We therefore also run a five variable VAR with the following ordering of variables: sentiment, the log real stock price index (as measured by the S&P 500), the three month t-bill rate, the log of real consumer spending, and the

to note that this procedure does not impose, or suggest, any particular structure on the data, but simply summarizes the historical dynamics between the variables included in the VAR. Hence, the VARs allow us to calculate the response, over time, of each spending category to a "representative" one-standard-deviation shock in the sentiment measure. The estimation strategy has the advantage of quantifying the impact of a change in consumer sentiment today on spending over time, so that we can compare the influence on spending tomorrow with that 10 quarters from now. Thus, we can trace out the dynamic consequences of an innovation in consumer attitudes. However, because such "impulse response" analysis takes into account how all the equations in the system interact, sentiment can be correlated with increases in spending through its influence on other variables such as (for example) labor income. The investigation simply shows the dynamic effect of changes in consumer attitudes on consumer spending, and unlike the reduced form analysis above, is silent about the *incremental* information content of each attitudinal measure.

Figures at the end of the text illustrate the dynamic response of consumer spending to changes in consumer confidence. In the impulse response functions, sentiment is allowed to influence consumer spending, labor income, and the financial indicators contemporaneously, but these variables influence sentiment only with a one period lag. This is consistent with the fact that survey results are known before spending and income data are released.²¹ Due to the large number of figures we limit our analysis to studying the dynamic impact of each index's broader measure of the consumer confidence or consumer sentiment.

log of real labor income. A Cholesky decomposition of the covariance matrix is used to obtain the responses. The variables are not first differenced even though some of them may be non-stationary because hypothesis tests based on the VAR in levels will have standard asymptotic distributions; see Sims, Stock, and Watson (1990), or Hamilton, (1994), chapter 20.

²⁰ This innovation is simply the portion of the change in sentiment that cannot be attributed to changes in the other variables in the system. Such a shift in attitudes could be thought of as changes in consumers' tastes arising from John Maynard Keynes's "animal spirits", or simply as shifts in preferences that the econometrician cannot measure.

²¹The fact that the financial controls are placed after the sentiment variables assumes that they cannot influence sentiment within the period. Ordering the financial indicators before the sentiment measure does not affect the results. Changing other orderings also does not affect the results.

Each figure shows four charts. The first chart shows the response (the percent increase) of the spending category to an increase in the sentiment measure in the full sample without the financial controls (that is, the VAR contains three equations, one for sentiment, spending and labor income). The second chart shows the response of the spending category to an increase in sentiment in the sample excluding the 1990-'91 recessionary period and without the financial controls; the third chart shows the response of the spending category in the full sample to an increase in sentiment including the financial controls, and the fourth chart shows the response of sentiment in the sample excluding the 1990-'91 recessionary period and including the financial controls.

The center line in each chart gives the impulse response, and the lines above and below the response are the 95th percentile probability bands. That is, there is a 95 percent probability that the true response lies within the outer bands. If the bands do not straddle the zero axis, there is a high probability that the true spending response to an increase in sentiment is not zero.²² In interpreting these figures, it is important to keep in mind that the results simply illuminate the timing of events between the variables in the system; they do not inform us about which measure of consumer attitudes may play a more "causal" role for consumer spending, nor do they reveal which measure has more incremental forecasting power.

We can compare the percent increase in total, personal consumption expenditure that results from an increase in the MI confidence index, with that resulting from an increase in the CB index by looking at Figures 6 and 11. The figures illustrate several interesting points. First, the MI index appears to have a more persistent effect on spending than the CB index does; after 10 quarters, total personal consumption expenditures are still about 0.5 percent above their pre-shock value using the MI index, whereas the affect on spending dies out after about three quarters using the CB index. Second, the magnitude of the standardized response is slightly larger using the MI index than using the CB index. For example, the south-east figure shows that an increase in the MI index leads to a 0.5 percent increase in total spending, whereas an increase in the CB index leads to a 0.3 percent increase. Third, the overall impact on spending of an increase in the MI index is substantially decreased once financial controls and the 1990-'91 recession are taken into account; the response

²²The bands are generated by taking 1000 Monte Carlo draws from the posterior distribution of the VAR coefficients.

of total, personal consumption expenditure is cut almost in half from about 1 percent to 0.5 percent as a result. In contrast, the reduction in that same response to a standardized increase in the CB index is much smaller, reducing the percent increase in total, personal consumption expenditures from 0.4 percent to 0.3 percent, approximately. This third observation confirms the existence of a high degree of colinearity between the MI index and the control variables already made apparent by the reduced form regressions results.

Other figures show a similar pattern of results. For example, Figures 7 and 12 show that increases in both indexes lead to a .20 percent increase in motor vehicle spending, once financial indicators and the 1990-'91 recession are controlled for. Again, shocks to the MI index die out more slowly, continuing to have a positive affect on spending up to 9 quarters later, while shocks to the CB index generally die out after about 3 quarters. A similar result holds for the impulse response of goods other than motor vehicles (Figures 8 and 13). Figures 9 and 14 show that the dynamic response of services to a shock in each index is generally very small and only marginally statistically significant. Finally, a comparison of Figures 10 and 15 demonstrates that the spending response of durables goods to a shock in the MI index does not reach its peak until many-quarters after the shock, while the peak reaction to a CB index impulse occurs almost immediately.

To summarize, the impulse response analysis illustrates at least two general points. First, the MI generally produces it's peak response several quarters later than the peak reaction induced by the CB index. Second, the influence on consumer spending of a shock to the MI index generally dies out after 5 or more quarters. In contrast, shocks to the CB index have a more immediate impact on spending, and their overall effect is more transitory.

Hints as to why impulses to the MI index appear to have more persistent effects on spending than do those to the CB index might again be found by analyzing the surveys' questions on consumer expectations. The MI survey asks consumers what they expect their situation to be one to five years from the interview date, whereas the CB survey asks consumers to forecast their circumstances only six months in the future.

4. Summary and Concluding Remarks

In this paper we have investigated the impact of consumer attitudes on consumer spending.

The main focus of our empirical analysis is to compare the forecasting power two measures of those attitudes. We also discuss the ways in which the surveys underlying these measures differ. The findings indicate that consumer attitudes have significant predictive power for several categories of consumer spending, even after economic fundamentals, and the particularly volatile movements in consumer attitudes associated with the Gulf War recessionary period are controlled for.

Perhaps more interesting, we find that these results are generally more robust for the Conference Board's measure of both overall confidence, and for its measure of consumer expectations, than they are for the University of Michigan's analogous measures. The Conference Board Index of Consumer Confidence appears to provide information about the future path of spending for several categories of consumption that is not captured by lagged values of labor income, stock prices, interest rates, or the spending category itself. Both the Michigan and Conference Board indexes have incremental forecasting power for expenditure on motor vehicles. These results contrast with those of other researchers (e.g., Carroll, Fuhrer and Wilcox, 1994, Fuhrer, 1993) who have focused their analysis on the Michigan Index of Consumer Sentiment; those studies find a very small role for consumer attitudes in predicting future movements in real consumption expenditures.

In addition to being a *statistically* significant predictor of real consumption expenditures, the CB index also appears to have *economically* significant effects; the increment to the fraction of regression variance explained from adding the CB index to the estimating equation is quite substantial for total, personal consumption expenditure, equal to 12 percent when using the CB expectations component. The findings should be of interest to academic economists because leading theoretical models of consumer behavior generally contain no role for sentiment, and they should be of interest to business economists because the forecasting power is quantitatively important.

We have left at least one important topic unexplored: the issue of what theoretical model might account for the spending-sentiment correlations we find. We caution that our evidence does not prove that consumer attitudes cause changes in consumer spending, though our reduced-form results are suggestive because they explicitly control for other economically fundamental factors that are thought to be important determinants of aggregate consumption growth. In short, the data on consumer confidence and consumer spending appear to contain at least some economically meaningful patterns.

Measures of consumer confidence are often used as a leading indicator of the state of the economy. Our empirical findings suggest that this function for sentiment is warranted, but demonstrate that the Conference Board measure is likely to be more useful in this regard for predicting the growth in a wider range of expenditure categories because it may contain information not captured by economic fundamentals. A remaining possibility is that we have not controlled for some other variable that is ultimately driving the sentiment-spending correlations found here. Nevertheless, the results indicate that household confidence may explain a significant fraction of the variance in consumption, and they raise the possibility that consumer attitudes act as a catalyst for economic fluctuations.

Data Appendix

Consumption

Real personal consumption expenditure (PCE), at quarterly frequency.

Five categories of the PCE are used - total, motor vehicles, goods excluding motor vehicles, services, and durables excluding motor vehicles. Source: the Bureau of Economic Analysis. The PCE for motor vehicles is calculated from the unit sales of new cars and trucks; PCE for other goods is estimated from the monthly retail sales data.

Labor Income

Labor income is defined as wages and salaries plus transfers minus personal contributions for social insurance. These components are quarterly from the National Income and Product Accounts (NIPA) data.

Interest rates

The interest rate used is the 3 month t-bill rate, available on a monthly basis from the Board of Governor's of the Federal Reserve System. The data are quarterly averages.

Stock Prices

Measured as the Standard and Poor's 500 composite stock price index (1941-43=10).

The stock price index is monthly and is available from the <u>Wall Street Journal</u> and the <u>Financial Times</u>.

Price Deflator

Nominal labor income and the S&P 500 are deflated by the PCE implicit price deflator (1992=100). The deflator is quarterly and taken from the National Income and Product Accounts (NIPA) data. The data reflects the revisions made by the Department of Commerce in September, 1993.

References

- Carroll, Christopher D., Jeffrey C. Fuhrer, and David W. Wilcox, 1994, "Does consumer sentiment forecast household spending? If so, why?", *American Economic Review*, 84(5), 1397-1408.
- Fair, R. C., 1971, "Consumer sentiment, the stock market and consumption functions," econometric Research Program Research Memorandum No. 119, Princeton University.
- Fuhrer, Jeffrey C., 1993, "What role does consumer sentiment play in the U.S. economy?", New England Economic Review, January/February, 32-44.
- Leeper, Eric M., 1992, "Consumer attitudes: king for a day", Economic Review, Federal Reserve Bank of Atlanta, July/August, 1-15.
- Ludvigson, Sydney C., 1996, "Consumer sentiment and household expenditure: reevaluating the forecasting equations", Federal Reserve Bank of New York Research Paper #9639.
- Mankiw, Gregory, 1982, "Hall's consumption hypothesis and durable goods", *Journal of Monetary Economics*, 10, 417-425.
- Matsusaka, John G., and Argia M. Sbordone, 1995, "Consumer confidence and economic fluctuations", *Economic Inquiry*, 33, April, 296-318.
- Mishkin, Frederic S., 1978, "Consumer sentiment and spending on durable goods," *Brookings Papers on Eonomic Activity*, 1, 217-232.
- Wilcox, David W., 1992, "The construction of U.S. consumption data: some facts and their implications for empirical work", *American Economic Review*, 82(4), 922-941.

Table 1

Reduced-Form Evidence: Adjusted R^2 's From Simple Prediction Equations $\Delta \log (C_t) = \alpha_0 + \sum_{i=1}^4 \beta_i S_{t-1} + \epsilon_t$

Overall Index

		Michigan Index		Conference Board Index	
Row	Category of Real PCE	Full <u>Sample</u>	Sample exc 90-91 rec.	Full <u>Sample</u>	Sample exc 90-91 rec.
1	Total	0.13 (0.000)	0.15 (0.000)	0.14 (0.002)	0.15 (0.007)
. 2	Motor Vehicles	0.17 (0.000)	0.17 (0.000)	0.16 (0.000)	0.16 (0.002)
3	Goods excl Motor Vehicles	0.17 (0.000)	0.18 (0.001)	0.15 (0.006)	0.15 (0.020)
4	Services	0.14 (0.002)	0.20 (0.001)	0.18 (0.001)	0.22 (0.001)
5	Durables Goods excl Motor Vehicles	0.13 (0.000)	0.13 (0.001)	0.04 (0.101)	0.04 (0.159)
Expec	tations Component				
1	Total	0.15 (0.000)	0.16 (0.000)	0.22 (0.000)	0.24 (0.000)
2	Motor Vehicles	0.18 (0.000)	0.18 (0.000)	0.24 (0.000)	0.24 (0.000)
3	Goods excl Motor Vehicles	0.17 (0.000)	0.18 (0.001)	0.17 (0.003)	0.18 (0.004)
4	Services	0.18 (0.000)	0.24 (0.000)	0.21 (0.000)	0.26 (0.000)
5	Durable Goods excl Motor Vehicles	0.10 (0.002)	0.10 (0.003)	0.04 (0.173)	0.04 (0.177)

Table 2

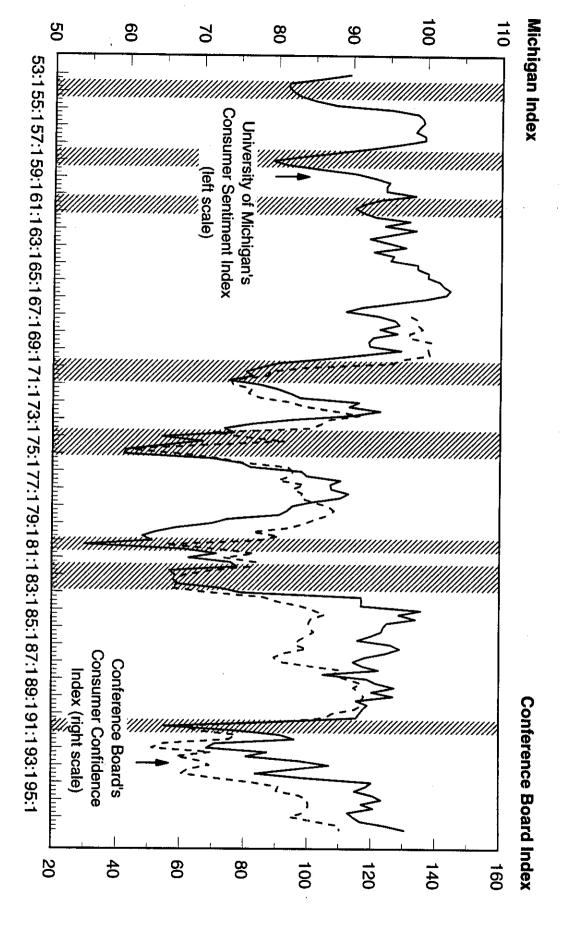
Reduced-Form Evidence: Incremental Adjusted R^2 's From Simple Prediction Equations $\Delta \log (C_t) = \alpha_0 + \sum_{i=1}^4 \beta_i S_{t,i} + \gamma \mathbf{Z}_t + \epsilon_t$

Overall Index

		Michigan Index		Conference Board Index	
Row	Category Real PCE	Full <u>Sample</u>	Sample exc 90-91 rec.	Full <u>Sample</u>	Sample exc 90-91 rec.
1	Total	0.07 (0.039)	0.01 (0.275)	0.11 (0.001)	0.09 (0.001)
2	Motor Vehicles	0.11 (0.011)	0.02 (0.059)	0.10 (0.004)	0.05 (0.020)
3	Goods excl Motor Vehicles	0.03 (0.135)	0.03 (0.084)	0.05 (0.013)	0.07 (0.177)
4	Services	-0.01 (0.698)	0.0 (0.434)	0.03 (0.075)	0.02 (0.062)
5	Durables Goods excl Motor Vehicles	0.02 (0.134)	0.01 (0.201)	0.14 (0.001)	0.15 (0.005)
Expec	tations Component				
. 1	Total	0.08 (0.029)	0.08 (0.194)	0.16 (0.000)	0.12 (0.000)
2	Motor Vehicles	0.07 (0.008)	0.02 (0.088)	0.13 (0.001)	0.10 (0.006)
3	Goods excl Motor Vehicles	0.01 (0.212)	0.08 (0.000)	-0.06 (0.056)	-0.12 (0.334)
4	Services	-0.01 (0.299)	0.02 (0.147)	0.02 (0.157)	0.06 (0.018)
5	Durable Goods excl Motor Vehicles	0.01 (0.215)	0.0 (0.272)	0.08 (0.142)	0.06 (0.217)

Notes: Sample: 1959:1-1996:3. The regressions underlying the adjusted R^{2} 's reported in Table 1 use only lagged values of consumer sentiment as independent variables. The regressions underlying the incremental adjusted R^{2} 's reported in Table 2 use the controls in addition to sentiment. These controls are lags 1 through 4 of the dependent variable, of growth in real labor income (defined as wages and salaries plus transfers minus personal contributions for social insurance), of growth in real stock price index, and of the first difference of the 3 month Treasury Bill rate. 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.

Figure 1
Consumer Attitudes



Quarters

FIGURE 2: PRESENT CONDITIONS Questions

Conference Board

- Q1) How would you rate present general business conditions in your area? [good / normal / bad]
- Q2) What would you say about available jobs in your area right now? [plentiful / not so many / hard to get]

Michigan

- Q1) Do you think now is a good or bad time for people to buy major household items?[good /neither /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 off / same / worse off]

Present Situation Component of Consumer Attitude Figure 3

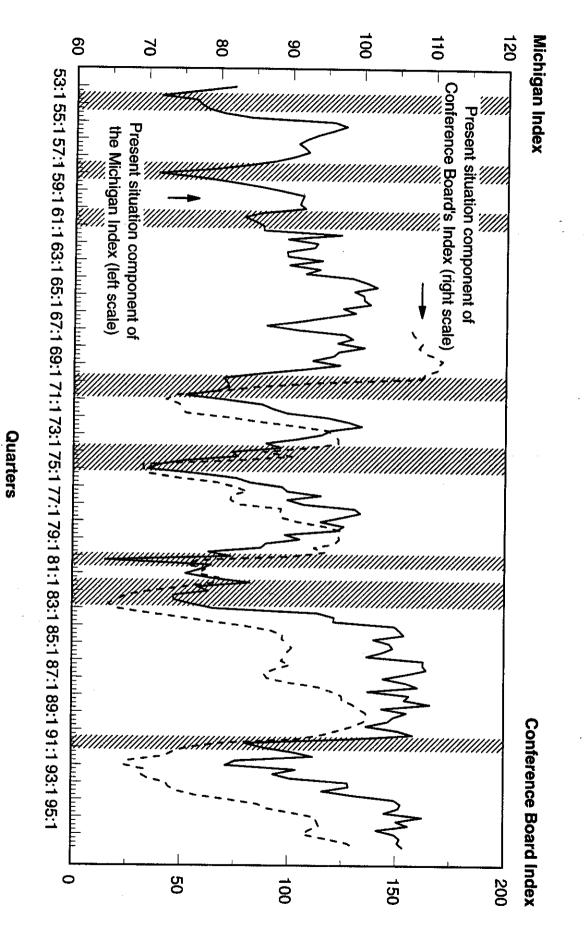


Figure 4: EXPECTATIONS Ouestions

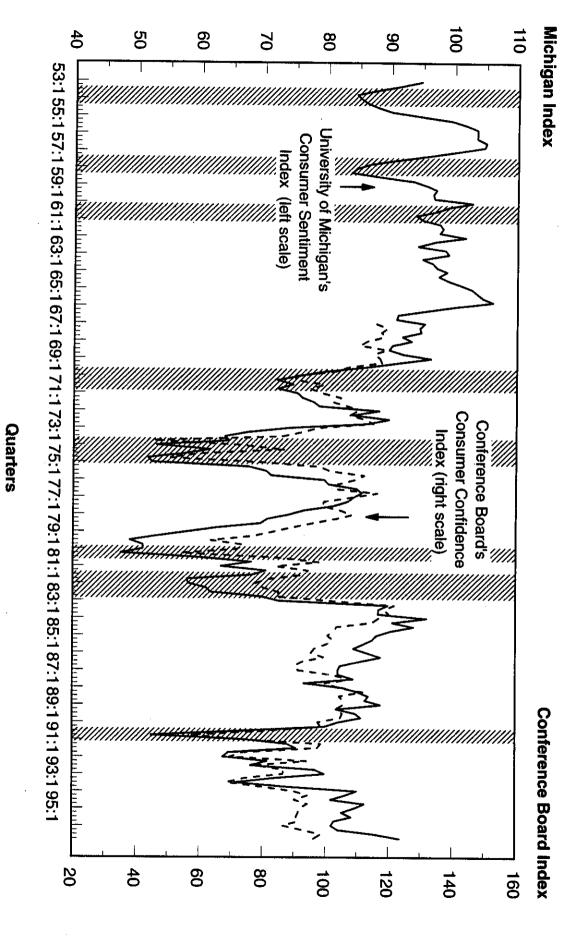
Conference Board

- Q3) Six months from now, do you think business conditions in your area will be [better / same / worse]?
- Q4) Six months from now, do you think there will be [more / same / fewer] jobs available in your area?
- Q5) How would you guess your total family income to be six months from now? [higher/same/lower]

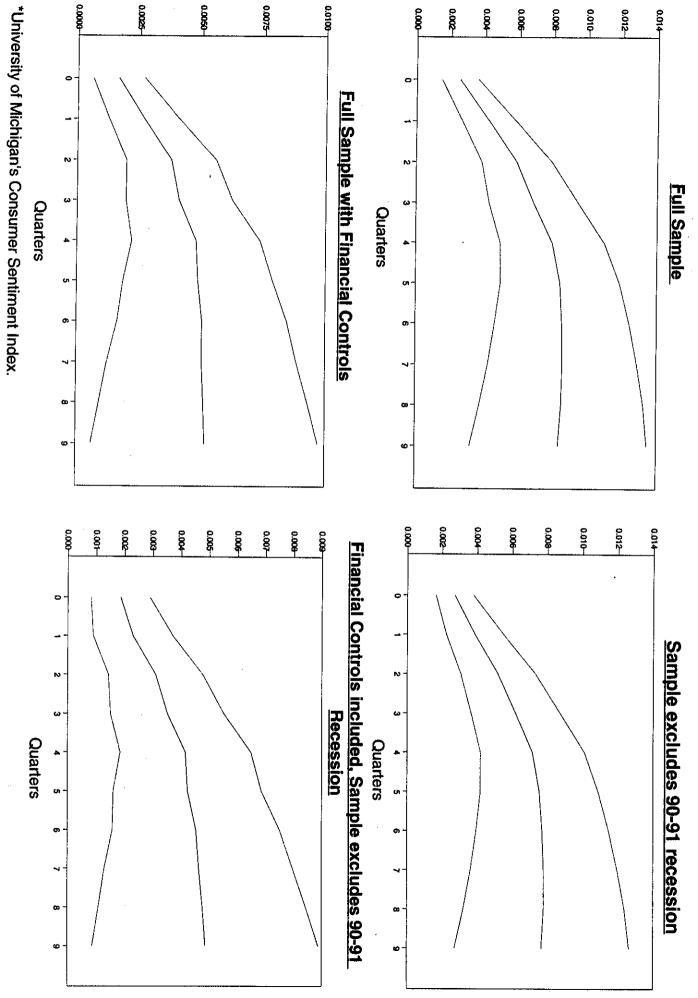
Michigan

- Q3) Now turning to business conditions in the country as a whole -- do you think that during the next 12 months, we'll have good times financially, or bad times or what? [good times / uncertain / bad times]
- 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 5 years or so, or that we'll have periods of widespread unemployment or depression, or what? [good times / uncertain / bad times]
- 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 off / same / worse off]?

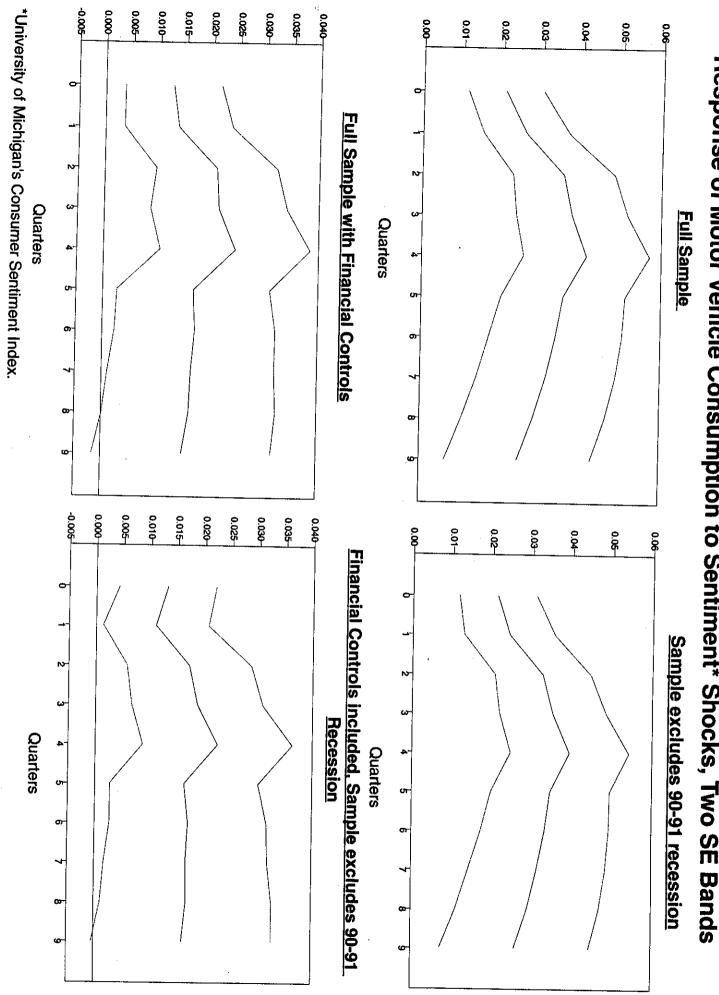
Figure 5
Consumer Expectations



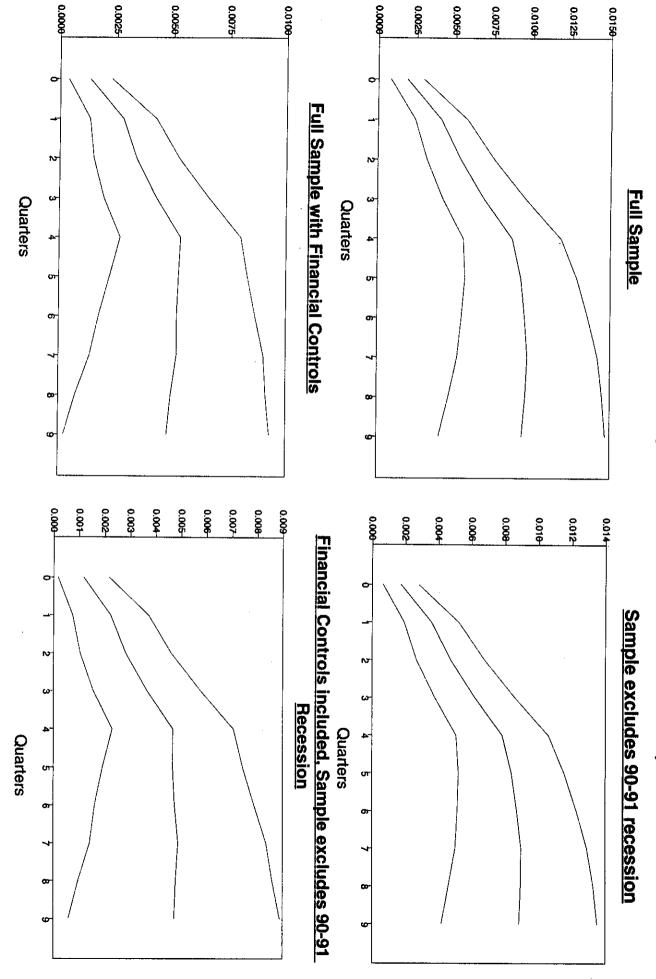
Response of Total Consumption to Sentiment* Shocks, Two SE Bands



Response of Motor Vehicle Consumption to Sentiment* Shocks, Two SE Bands Figure 7



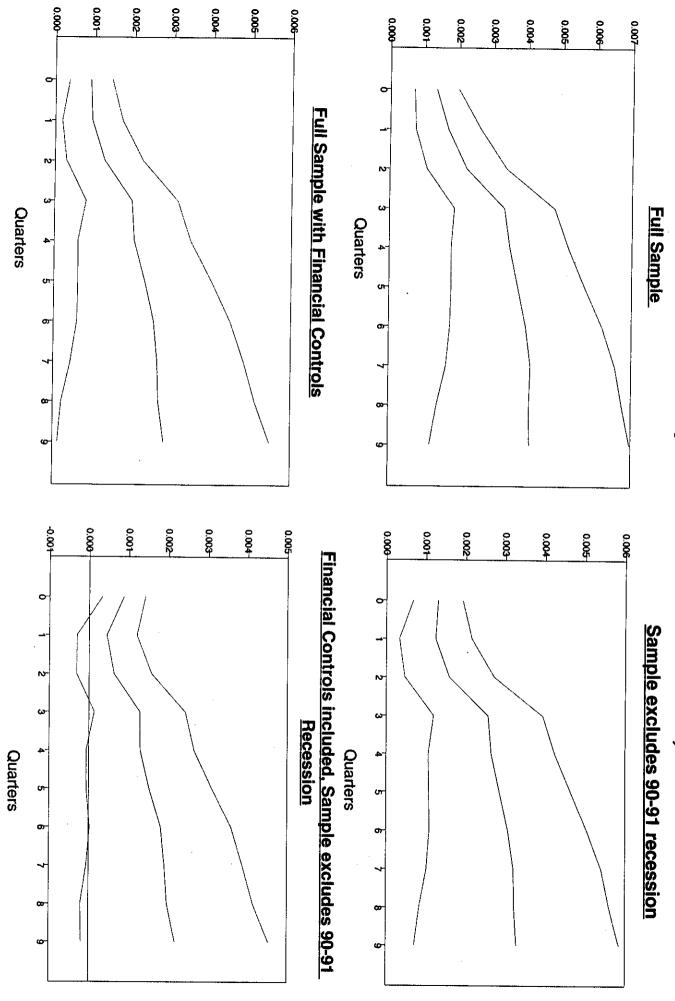
Response of Other Goods* Consumption to Sentiment** Shocks, Two SE Bands



^{*} Durable and non-durable goods excluding motor vehicles.

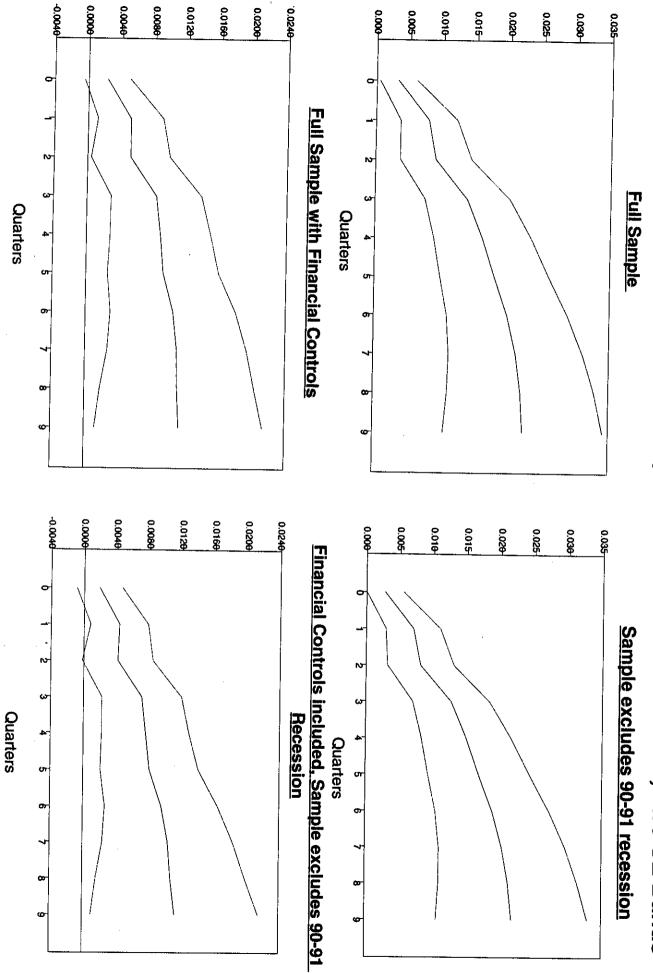
^{**}University of Michigan's Consumer Sentiment Index.

Response of Services Consumption to Sentiment* Shocks, Two SE Bands Figure 9



*Úniversity of Michigan's Consumer Sentiment Index.

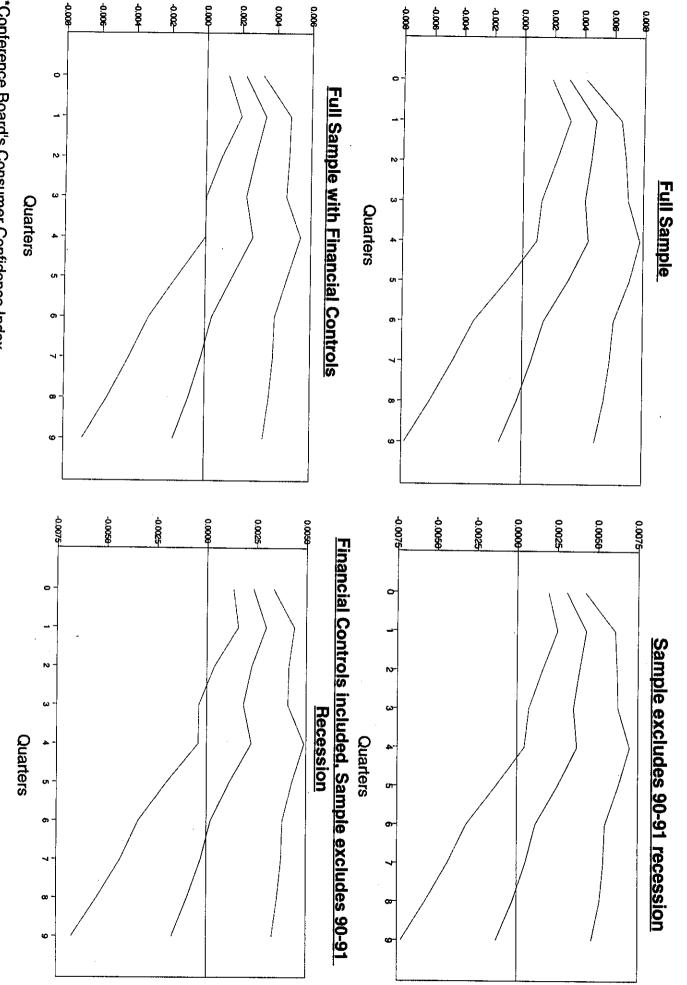
Response of Durable Goods* Consumption to Sentiment** Shocks, Two SE Bands



^{*} Durable goods excluding motor vehicles.

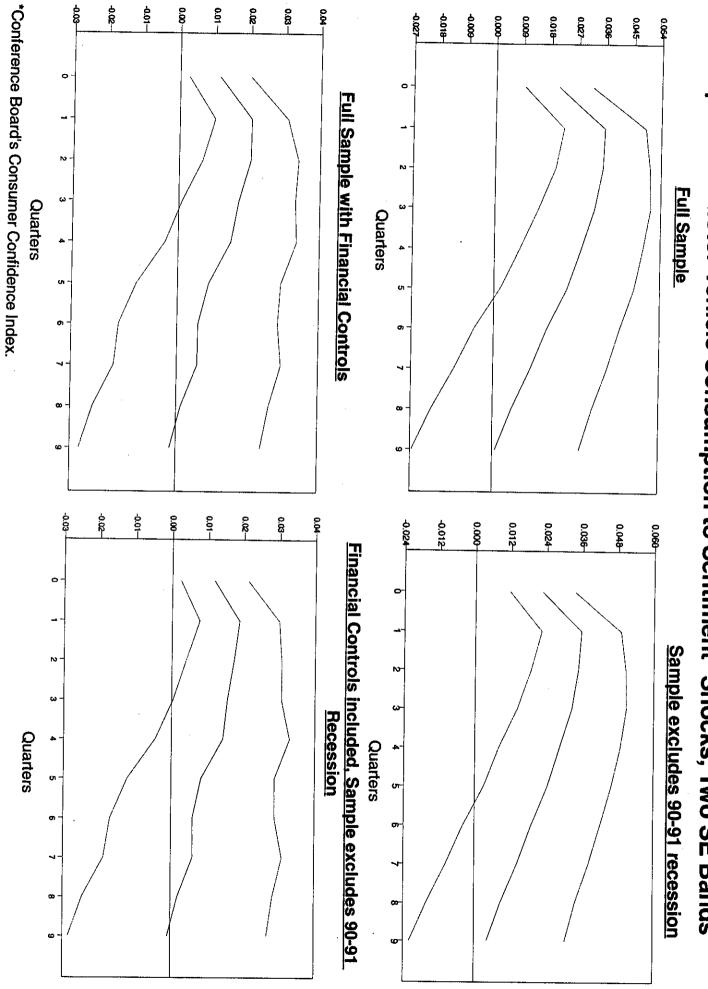
^{**}University of Michigan's Consumer Sentiment Index.

Response of Total Consumption to Sentiment* Shocks, Two SE Bands

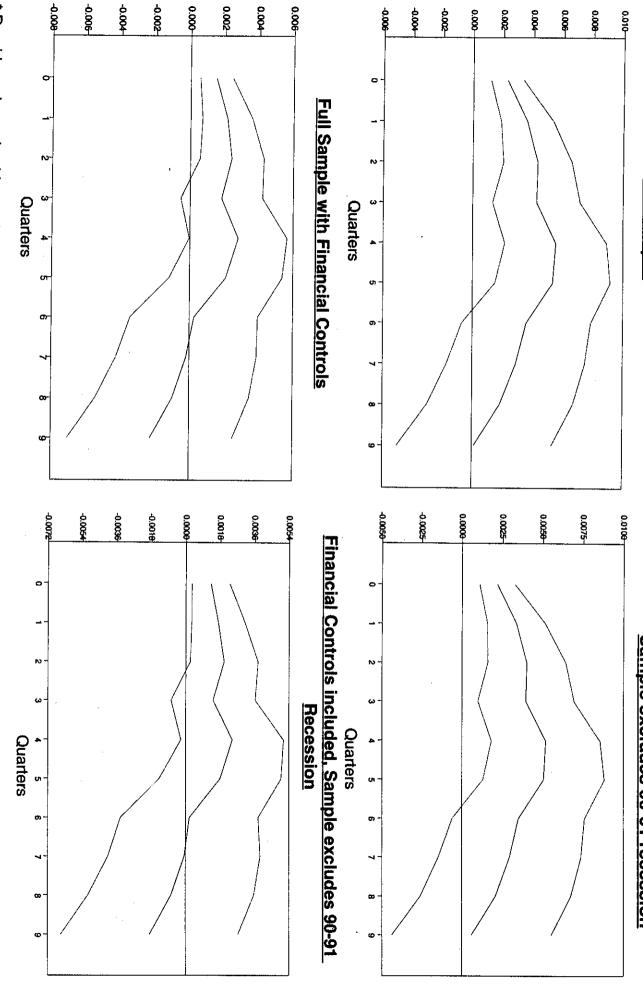


*Conference Board's Consumer Confidence Index.

Response of Motor Vehicle Consumption to Sentiment* Shocks, Two SE Bands



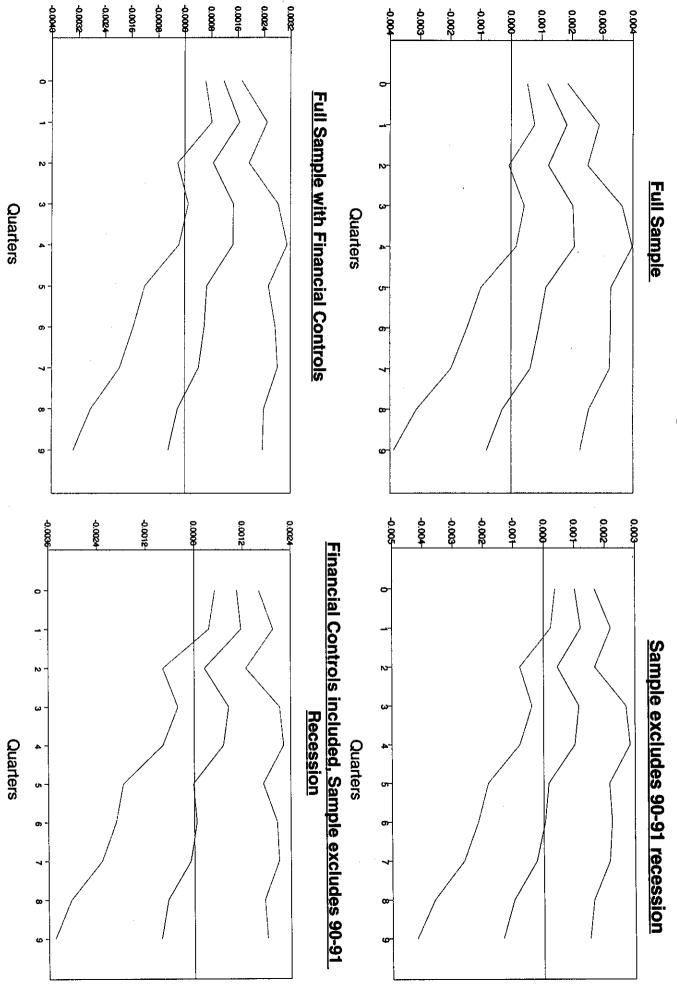
Response of Other Goods* Consumption to Sentiment** Shocks, Two SE Bands Full Sample Figure 13 Sample excludes 90-91 recession



^{*} Durable and non-durable goods excluding motor vehicles.

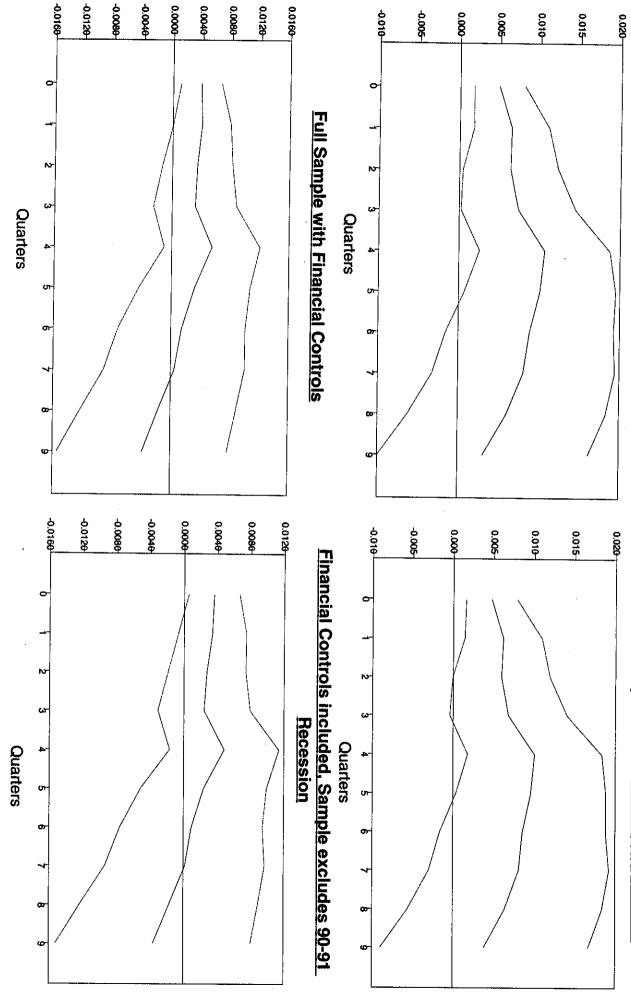
^{**}Conference Board's Consumer Confidence Index.

Response of Services Consumption to Sentiment* Shocks, Two SE Bands Figure 14



^{*}Conference Board's Consumer Confidence Index.

Response of Durable Goods* Consumption to Sentiment** Shocks, Two SE Bands Full Sample Figure 15 Sample excludes 90-91 recession



^{*} Durable goods excluding motor vehicles.

^{**}Conference Board's Consumer Confidence Index.

FEDERAL RESERVE BANK OF NEW YORK RESEARCH PAPERS 1997

The following papers were written by economists at the Federal Reserve Bank of. New York either alone or in collaboration with outside economists. Single copies of up to six papers are available upon request from the Public Information Department, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045-0001 (212) 720-6134.

- 9701. Chakravarty, Sugato, and Asani Sarkar. "Traders' Broker Choice, Market Liquidity, and Market Structure." January 1997.
- 9702. Park, Sangkyun. "Option Value of Credit Lines as an Explanation of High Credit Card Rates." February 1997.
- 9703. Antzoulatos, Angelos. "On the Determinants and Resilience of Bond Flows to LDCs, 1990 1995: Evidence from Argentina, Brazil, and Mexico." February 1997.
- '9704. Higgins, Matthew, and Carol Osler. "Asset Market Hangovers and Economic Growth." February 1997.
- 9705. Chakravarty, Sugato, and Asani Sarkar. "Can Competition between Brokers Mitigate Agency Conflicts with Their Customers?" February 1997.
- 9706. Fleming, Michael, and Eli Remolona. "What Moves the Bond Market?" February 1997.
- 9707. Laubach, Thomas, and Adam Posen. "Disciplined Discretion: The German and Swiss Monetary Targeting Frameworks in Operation." March 1997.

To obtain more information about the Bank's Research Papers series and other publications and papers, visit our site on the World Wide Web (http://www.ny.frb.org). From the research publications page, you can view abstracts for Research Papers and Staff Reports and order the full-length, hard copy versions of them electronically. Interested readers can also view, download, and print any edition in the Current Issues in Economics and Finance series, as well as articles from the Economic Policy Review.