CONSUMER CONFIDENCE AND ECONOMIC FLUCTUATIONS

JOHN G. MATSUSAKA and ARGIA M. SBORDONE*

If consumers become pessimistic about the state of the economy, can there be a slowdown in output, even if their pessimism is not based on economic fundamentals? Recent macroeconomic models show the answer is yes, if there are "strategic complementarities" and multiple equilibria. We investigate the link between consumer confidence and economic fluctuations using vector autoregressions. In all models, after controlling for economic fundamentals, the hypothesis that consumer sentiment does not cause GNP (in the Granger sense) can be rejected. Variance decompositions suggest that consumer sentiment accounts for between 13 and 26 percent of the innovation variance of GNP.

I. INTRODUCTION

The idea that consumer sentiment might cause output fluctuations is popular in the business press.[1] It is easy to see where this belief comes from: in Figure 1 we plot an index of consumer sentiment against recessions for the postwar period.[2] The pattern is striking; all recessions were preceded by a fall in confidence, and all major falls in consumer sentiment were followed by a recession (except in 1965 which, while not a recession, was the so-called "growth recession"). Apparently either consumers were correctly forecasting output falls, or declines in consumer sentiment were inducing declines in output.

It has been recognized for some time that business cycle and growth models can exhibit multiple (Nash) equilibria if they contain strategic complementarities.[3] In a conventional single-equilibrium model, output fluctuates only in response to changes in economic fundamentals, where the list of fundamentals includes technology, government purchases, money supply, the price of oil, and so on. In a multiple-equilibria model output responds to fundamentals, but in addition there can be fluctuations as the economy shifts between equilibria. Perhaps the most intriguing feature of these models is

* Hoover Institution, Stanford University, Stanford CA 94305-6010 and Department of Finance and Business Economics, School of Business Administration, University of Southern California, Los Angeles CA 90089-1421; and Economic Research Department, Federal Reserve Bank of Chicago, Chicago, Illinois, 60604-1413. We thank Joonmo Cho, Giorgio De Santis, Robert Erikson, Selahattin Imrohoroglu, Douglas Joines, Narayana Kocherlakota, Robert Kollman, Fulvio Ortu, Ian Parry, Christopher Phelan, Pedro Teles, Mark Watson, Michael Woodford, anonymous referees, and workshop participants at the University of Chicago for many helpful comments.

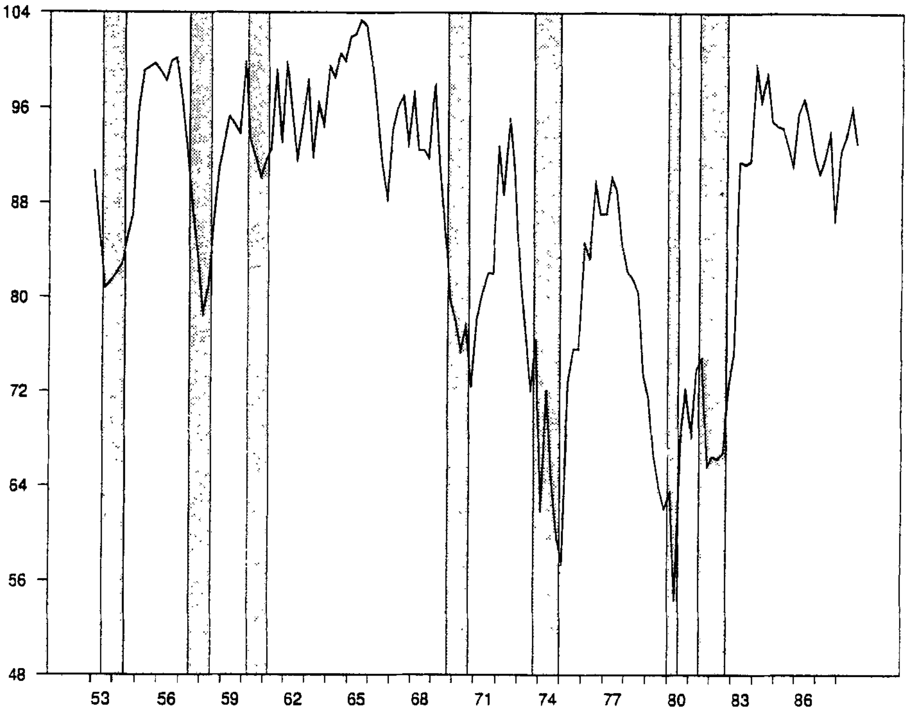
1. Examples from the last recession are "Economy in the U.S. Isn't Nearly as Sour As the Country's Mood. But Pessimism Could Become A Self-Fulfilling Prophecy Further Stalling Recovery. Can Attitude Be Everything?" on the front page of *The Wall Street Journal*, November 4, 1991, and "Recession: Case of Bad Attitude? President Bush says people should be more confident—and that would restore economic vigor. But consumers cite layoffs and lost wages," from the front page of the *Los Angeles Times*, November 9, 1991.

2. A recession as defined by the National Bureau of Economic Research is indicated with a vertical bar.

3. Roughly speaking, a model contains strategic complementarities if each agent's optimal action is positively correlated with the action of other agents. A representative but incomplete list of the literature would include Ball and Romer [1991], Bryant [1983; 1987], Cass and Shell [1983], Kiyotaki [1988], Milgrom and Roberts [1990], Murphy, Shleifer, and Vishny [1989], Roberts [1987], Shleifer [1986], and Weil [1989]. Cooper and John [1988] provide a good overview. On theoretical objections see Woodford [1987].

Economic Inquiry
(ISSN 0095-2583)
Vol. XXXIII, April 1995, 296–318

FIGURE 1
Index of Consumer Sentiment (1953–1988)



that output can fluctuate simply because everyone expects it to. Put differently, expectations can be self-fulfilling in that if people expect bad times they get them. Thus, this class of models offers a theoretical rationalization of how consumer sentiment can affect GNP. Although these models introduce dramatically different policy considerations than single-equilibrium models, many of their observable implications are the same. As a result, their empirical relevance is largely unknown.

In this paper, we study the behavior of the U.S. economy for the period 1953–1988. Our central purpose is to evaluate empirically how much truth there is, if any, to the idea that consumer sentiment causes fluctuations in GNP. Following Granger and others we take causality to imply temporal precedence, that is, if one series causes another, then movements in the forcing series precede movements in the forced series. We verify that the temporal ordering suggested in Figure 1, consumer sentiment leads GNP, is statistically significant. Temporal precedence is not enough for our purposes, however, because the evidence that movements in consumer confidence precede movements in output can be interpreted in two ways. Either consumer sentiment *causes* GNP or it simply *anticipates* GNP.

We attempt to disentangle the two possibilities by estimating a series of vector autoregressions including GNP, consumer sentiment, and various other series that capture fundamentals or are good predictors of GNP. The working assumption is that if consumer sentiment can forecast GNP movements even after controlling for fundamentals and other publicly available predictors of output, some support is provided for the idea that consumer sentiment causes output fluctuations. In our baseline regressions, the control variable is an index of leading indicators, composed of fundamentals and purely predictive variables. To check for robustness we consider a number of other controls, a "textbook" set of fundamentals that includes government spending, money supply, and sensitive materials prices; components of the leading indicators; and a default risk variable that has been identified as an excellent predictor of output movements. For each set of vector autoregressions we test whether consumer sentiment Granger-causes GNP.

Our main finding is that even after controlling for economic fundamentals and other good predictors of GNP, changes in consumer sentiment have a statistically significant effect on output fluctuations. In other words, we find evidence of Granger causality running from consumer sentiment to GNP. We also attempt to assess the economic significance of consumer sentiment by decomposing the variance of GNP innovations. Our second finding is that while sentiment is not the most important factor in GNP fluctuations, it plays a quantitatively significant role: between 13 percent and 26 percent of GNP innovation variance can be attributed to innovations in consumer sentiment. The evidence suggests that consumer confidence is an important independent factor in economic fluctuations, and more tentatively, it appears to provide some support for the class of macroeconomic models with strategic complementarities.

The paper proceeds as follows. The next section provides a simple illustrative model that exhibits multiple equilibria for some parameter values. Its main features are consistent with the bulk of the macroeconomics literature on strategic complementarities, but we hope it is more transparent. The model also calls attention to some interesting features of actual economies that have been underexplored in this context. Section III discusses the related literature. In section IV we present the evidence. Granger causality tests are reported that assess how important, if at all, consumer sentiment is in determining the level of aggregate production. Section V draws a few conclusions.

II. AN ILLUSTRATIVE MODEL

It is easier to discuss the empirical procedures with a concrete example in mind. In this section we sketch a stylized model of output fluctuations that exhibits the basic features of multiple-equilibria models. The model is amenable to generalization in a number of directions that we do not pursue; our goal is to illustrate the central issues in as simple a manner as possible.

The model is motivated by two features of actual economies that have not received much attention in the recent macroeconomics literature: (i) many goods are produced to buyer specifications, and (ii) once goods are customized they become less valuable to other buyers. In fact, production to order is important in such industries as construction, machine tools, steel, women's apparel, and durables manufactures in general (Zarnowitz [1973]). What makes this interesting is that these are the industries where the business cycle primarily occurs.

We model customization by assuming there is one consumption good, call it a house, that a person can have made to her specifications—she can choose the color of its paint. She can also buy a house on the resale market or from developers who build without having a buyer lined up. Because a resold house and a developer's house are not specialized to her tastes—

the paint is the wrong color—they give her lower utility. In our one-period model consumers submit orders for custom-made goods at the start of the period, keeping their expected income in mind. If they order more than they can afford, they must resell part of their order, which is costly because other buyers place a relatively low value on the custom features. If they order too little, they must enter the resale market as buyers, which means they must buy goods whose custom features they dislike. Because consumers strive to match their orders with their income, expectations can be self-fulfilling: if, for some reason, people expect low incomes they submit low orders, leading the build-to-order firms to have low employment, and consumers to have low incomes.

We assume there is a continuum of identical agents over the interval [0,1]. Agent i has a utility function of $u(i) = c(i) - \frac{1}{2} l(i)^2$ where $l(i)$ is her labor and $c(i)$ is her consumption value. To capture the idea of specialized goods, her consumption value is written $c(i) = c_0(i) + \tau c_1(i)$ where $c_0(i)$ is goods she has custom made and $c_1(i)$ is goods that are not custom made. She prefers goods that are customized: $\tau \leq 1$.

There is a competitive production sector that produces Y goods where

$$Y = A \int_0^1 l(i) di.$$

All integrals that follow are over [0,1]. Firms are jointly and equally owned and all profits are immediately returned as dividends (although in equilibrium profits are zero). The aggregate production function (we sometimes call it "the firm") is completely described by A; this parameter represents economic fundamentals. The production technology is the same for custom goods and generic goods so $\int c_0(i) di + \int c_1(i) di = Y$. One way to think of this is that the cost of painting a house is the same for every color. If labor demand exceeds labor supply we assume that employment is rationed equally among workers.[4]

The sequence of actions in a period is as follows. First, each agent i submits orders for a quantity $q(i)$ of custom-made goods taking as given the price of custom-made and generic goods. This is the upper bound on person i's consumption of custom goods, that is, $c_0(i) \leq q(i)$. After orders are placed the firms produce and pay wages. Then workers are obligated to take delivery of their orders and a secondary market for consumption goods opens, which we call a resale market. Workers who do not have sufficient funds to pay for their orders must enter the resale market and sell part of their goods to pay off their debts. In addition, firms can sell whatever non-customized goods they produced. When trading is concluded consumption occurs.

Let p_0 be the price of custom-made goods, p_1 be the price of goods in the resale market, and w be the wage rate. We normalize $p_0 = 1$ and think of the wages paid by the firm as claims to a unit of the consumption good. These claims can be redeemed for one unit of the good; this puts a ceiling on the equilibrium price in the resale market: $p_1 \leq 1$.

In the resale period, if there are consumers who ordered more than they can afford, they attempt to raise money by selling some of their orders. They will be met by people who ordered less than they can afford and are trying to spend their extra income (claims). The interactions of these buyers and sellers determines p_1. It turns out that there is a rational expecta-

4. One way this can be rationalized is by assuming that each worker produces a separate component of the consumption good. A more formal way to model this behavior would be to write the production technology as $Y = A \min_i l(i)$. In keeping with the spirit of our model as an example, we eschew this additional formality because it makes the same point while introducing a heavier burden of notation.

tions equilibrium for any $p_1 \in [0,1]$.[5] The rest of this example is confined to the case where $p_1 = \tau$. This is an appealing case because goods in the resale market are discounted just enough to compensate for the lower utility they deliver.

The next two propositions show that the behavior of the economy depends in an important way on the value of τ. In particular, when $\tau = 1$ there is a unique equilibrium where output is determined only by fundamentals. When $\tau < 1$ there is a continuum of equilibria and expectations have self-fulfilling effects on the level of output. Here and below an asterisk is used to denote equilibrium quantities, for example, equilibrium utility for agent i is $u*(i) = c*(i) - ½\, l*(i)^2$.

PROPOSITION 1: *If $\tau = 1$ there is a unique equilibrium with $Y = A^2$.*

Proof. If $\tau = p_1 = 1$ the firm finds it equally profitable to sell in the custom market and the resale market. Then the firm's profit is $A \int l(i)di \; - w \int l(i)di$. Equilibrium requires $w* = A$ and the firm hires as many as will work at that wage. From the consumer's maximization problem, $l*(i) \; = w* \; = A$. Then $Y* = A \int l*(i)di = A^2$. Q.E.D.

The intuition behind the proposition is that when $p_1 = \tau = 1$ the price in the ordered goods market is the same as the price in the resale market, so from the firm's point of view a good produced to order is as valuable as a good produced to stock. Consequently, the firm pays a wage equal to A and hires as much labor as is supplied at that wage. Given the utility function, each person supplies A units of labor. The equilibrium is Pareto optimal.

In the situation described by Proposition 1 there is a unique equilibrium. There is no sense in which consumer expectations affect production. Output fluctuates only in response to changes in economic fundamentals, A. In these situations, when there are no customizing effects, the model behaves in what might be called the conventional way.

The next proposition states that when $\tau < 1$ there is a continuum of equilibria. The argument has two parts. First, the firm produces only to order, that is, it produces only as many goods as are ordered in the customized market. The reason is that if the firm produces for the resale market the wage rate is p_1A, and after working enough to supply the custom-made demand, agents will not supply additional labor at the wage p_1A. Second, given their income, which they take as exogenous, consumers order exactly as many customized goods as they can pay for. They lose if they enter the resale market either as buyers or sellers. In these cases a coordination failure occurs; if everyone could get together and jointly submit high orders efficiency would be attained.

PROPOSITION 2. *If $\tau < 1$ there is an equilibrium with $Y = SA^2$ for any $S \in (\tau, 1)$.*

Proof. Consider an arbitrary $S \in (\tau,1)$. The proof demonstrates that there is an equilibrium for S where $q^*(i) = SA$, $l^*(i) = SA$, and $w* = A$. In the proposed equilibrium person i is willing to supply labor of $l(i) = A$ but the demand is only $l^*(i) = SA$, according to the rationing rule. At the margin if the firm adds a unit of labor it generates revenue p_1A. But the marginal disutility of labor at $l^*(i)$ is $SA > p_1 A$.

Agents are atomistic so they take $l^*(i)$ and $w*$ as given. If person i orders $q(i) > q*(i)$ then she has to raise $q(i) - q*(i)$ in the resale market to cover her debts. This costs her $[q(i) - q*(i)]/p_1$ in consumption. Her utility is then

5. We shall not prove this here, but note that it follows from the fact that no equilibrium has an active market in both customized and non-customized goods. In practice, exactly where the resale price settles will depend on the structure of the resale market and the nature of the bargaining that takes place.

$$u(i) = q(i) - [q(i) - q*(i)]/p_1 - \frac{1}{2}l*(i)^2$$
$$= q*(i) - \frac{1}{2}l*(i)^2 + (1 - 1/p_1)[q(i) - q*(i)]$$
$$< u*(i),$$

where the inequality follows because $p_1 = \tau < 1$. If she orders $q(i) < q*(i)$, then she has $w*l*(i) - q(i) = q*(i) - q(i)$ to spend in the resale market. This buys $c_1(i) = [q*(i) - q(i)]/p_1$. Her utility is then $u(i) = q(i) + \tau \quad [q*(i) - q(i)] \quad /p_1 - \frac{1}{2}l*(i)^2 = q*(i) - \frac{1}{2}l*(i)^2 = u*(i)$. Thus, person i has no incentive to place higher or lower orders. *Q.E.D.*

The proposition implies that the level of output is determined not only by economic fundamentals, A, but also by an arbitrary parameter, S, that can be thought of as consumer sentiment. In an important sense it can be said that consumer expectations *cause* output. If consumers believe the equilibrium will be S then it is individually rational for them to take actions that bring about the S equilibrium. The interesting conclusion is that aggregate production can fluctuate in response to waves of optimism and pessimism, even if the underlying technology does not change. Perhaps more intriguing, in view of the recent behavior of the U.S. economy, is the suggestion that GNP can become stalled at low levels even when there is nothing wrong with the fundamentals of the economy.

III. RELATED LITERATURE

How can we tell if strategic complementarities are a real concern, as the business press believes, or merely of theoretical interest? One approach would be to try to measure τ or $p_0 - p_1$. Intuition suggests it is easy for parameter values to lie in the necessary range. For example, suppose a person orders a house built. She has many options, what color paint, what fixtures, how many rooms, what kind of carpeting or floors, and so on. If she has to sell the house, probably she will need to give the buyer a price break because her custom features are not the buyer's first choice. It is plausible, then, that $p_1 < p_0$ for durable goods; if $\tau < 1$ there are multiple equilibria. An obvious problem with this way of examining the theory is that it is intimately linked to our model, which is quite stylized.

Still, this example suggests that one way to test for multiple equilibria is to estimate directly the parameters of the model. Dagsvik and Jovanovic [1991] is a nice example of this approach. They estimate a fairly general Phillips curve capable of exhibiting multiple equilibria; then the parameters are examined to see if they lie in the multiple equilibria region of the parameter space. Their evidence indicates unique equilibria in all three time periods they investigate: 1901–1940, 1921–1940, and 1951–1986. The authors note that their approach is limited in that they are forced to posit an exogenous and constant equilibrium selection mechanism.[6]

Oh and Waldman [1990; 1993] take a completely different approach, one that is closer in spirit to ours. Rather than try to estimate a particular model, they identify a general characteristic of coordination failure models and look for behavior consistent with that feature. Cooper and John [1988] show that what drives the multiple equilibria property of all these models is "strategic complementarities": if everyone but person i takes the "high" action (in our model, submits high orders) then it is in person i's interest to take the high action as well.

6. Imrohoroglu [1993] uses a similar approach to look for non-fundamentals in the 1923 German hyperinflation, with largely negative results.

Oh and Waldman observe that when there are strategic complementarities people have an incentive to try to catch the waves: they want to invest when output is high and cut back when it is low. Therefore, in a world with strategic complementarities investment and output should respond to what people expect output to be. Their key insight is that if there is an announcement that the economy is about to boom and everyone believes it then future output should be high, even if the announcement is based on false information. In effect, people use the announcements to coordinate.

As a measure of false forecasts, Oh and Waldman use revisions of the government's *Index of Leading Indicators.* They look to see if these revisions had an effect on future economic activity. Their evidence is mixed, but more often than not it supports the idea that index revisions predict movements in industrial production, consistent with strategic complementarities. However, as they note, if the *Index of Leading Indicators* provides new information to firms and consumers (that is, information other than what equilibrium is selected), then revisions can have effects even when there are no strategic complementarities.

In this paper we focus on another implication of multiple-equilibria models: expectations are self-fulfilling. Specifically, we estimate a number of vector autoregressions and ask whether consumer sentiment about the future Granger-causes GNP.[7] The basic intuition behind the test comes from Proposition 2, which shows that the level of output is $Y = SA^2$. In contrast to models with a unique equilibrium, in a multiple-equilibria model consumer sentiment (S) has an effect on output (Y) even after controlling for fundamentals (A). All coordination failure models generate an output relation analogous to Proposition 2 because consumers must expect to be in a particular equilibrium in order for the economy to move there.

IV. EVIDENCE

Description of the Data and Variables

To measure consumer sentiment, we use the *Index of Consumer Sentiment* constructed by the Survey Research Center at the University of Michigan. This index captures consumer confidence about the economic present and future from survey responses to the five questions listed in Panel A of Table I.[8] The output series is GNP. We also performed the tests substituting the value of goods output for GNP, and arrived at similar conclusions. The main control variable is the *Index of Leading Indicators* (*Leading Indicators*), compiled by the Bureau of Economic Analysis at the U.S. Department of Commerce. We use the version of the *Leading Indicators* that expired in 1989–after 1989, the *Leading Indicators* includes the *Index of Consumer Expectations,* which is based on a subset of the questions used to calculate the *Index of Consumer Sentiment* (Hertzberg and Beckman [1989]). The series comprising the *Leading Indicators* are indicated in Panel B of Table I. In various regressions we also include as controls government expenditure, the components of the *Leading Indicators,* and the difference between the rate of interest on six-month commercial paper and the rate of interest on six-month Treasury bills, which we call "default risk."

7. Our approach is related to an older literature that investigated the ability of consumer attitudes to predict purchases of durable goods, for example, Friend and Adams [1964].

8. The Survey Research Center publishes a second index called the Index of Consumer Expectations, based on the answers to questions 2, 3, and 4. We use the *Index of Consumer Sentiment* instead of the *Index of Consumer Expectations* for two reasons. First, we believe the responses to questions 1 and 5 give a fuller picture of consumer sentiment. Second, we found that the *Index of Consumer Expectations* gave unstable results across subsamples. See Curtin [1982] for a more extensive discussion of the consumer expectations index.

TABLE I
Components of the Index of Consumer Sentiment and Index of Leading Indicators

A. Index of Consumer Sentiment

1. We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago? Why do you say so?
2. 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?
3. 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?
4. 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?
5. About the big things people buy for their homes—such as furniture, a refrigerator, a stove, television, and things like that. Generally speaking, do you think now is a good or a bad time for people to buy major household items? Why do you say so?

B. Index of Leading Indicators (1988 Version)

1. Average weekly hours of production or nonsupervisory workers in manufacturing.
2. Average weekly initial claims for unemployment insurance in state programs.
3. Manufacturers' new orders in consumer goods and materials industries.
4. Contracts and orders for plant and equipment.
5. Index of new private housing units authorized by local building permits.
6. Index of stock prices of 500 common stocks.
7. Money supply, M2.
8. Percent of companies receiving slower deliveries from vendors.
9. Change in sensitive materials prices.
10. Change in business and consumer credit outstanding.
11. Change in manufacturing and trade inventories on hand and on order.

The data run quarterly from 1953:I to 1988:IV, except for the default risk variable which runs from 1959:I. GNP and the *Leading Indicators* are seasonally adjusted at the source; the *Index of Consumer Sentiment* does not appear to have a seasonal component. GNP and the *Leading Indicators* are expressed in log differences. This transformation makes the series approximately mean and variance stationary.

The questions comprising the *Index of Consumer Sentiment* ask about changes as well as about levels, suggesting that the index is a measure of change in sentiment. Given this and the fact that the index is trendless, the estimates use the consumer sentiment data in levels. Default risk is also expressed in levels. Rather than take a strong stand on a particular number of lags, in most cases we report estimates from models with lags of one, two, three, and four quarters, although we find that the four-lags model tends to provide the best fit. The effects under consideration are weak with more than four lags. Ten observations are missing from the *Index of Consumer Sentiment* for the 1950s and one for the 1960s. This makes for 133 observations. The data are published by the U.S. Department of Commerce in *Survey of Current Business*.

Output and Consumer Sentiment

Causality is a difficult concept in science, and particularly so in economics where we seldom have the luxury of controlled experiments. The most common empirical approaches in economics to causality inference are based on the idea of "Granger causality." Roughly speaking, series x is said to Granger-cause series y if movements in x help to predict movements in y. Granger causality has the look of "true" causality so we believe it provides a natural stepping off point for empirical investigation. However, it should be kept in mind that Granger causality is neither a necessary or sufficient condition for true causality to exist.[9]

Our empirical strategy is to demonstrate first that Granger causality exists from consumer sentiment to GNP, and second, that this relation is robust to inclusion of a large number of alternative specifications. We begin with a simple two-variable system including only GNP and the consumer sentiment index. Two types of evidence are reported to test if changes in consumer sentiment predict changes in output. The first test is the so-called "Granger test." Bivariate vector autoregressions are estimated of the form

$$\begin{bmatrix} Y_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} Y_t \\ S_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{Yt} \\ \varepsilon_{St} \end{bmatrix}$$

where Y_t is GNP growth in quarter t, S_t is the *Index of Consumer Sentiment*, $x(L)$ is a polynomial in the lag operator L, and ε is an error.[10] If the block of coefficients represented by $b(L)$ are not jointly equal to zero, then we can reject the hypothesis that there is not Granger causality from consumer sentiment to GNP. The system also provides a statistical summary of the link apparent in Figure 1.

Table II presents the results for one, two, three, and four lags.[11] In the table, each column is a regression. The number of lags and the dependent variable are indicated at the top of each column. The main entry for each variable is the p-value for the F-statistic testing whether the block of coefficients are jointly equal to zero. In all four models, the hypothesis that the consumer sentiment coefficients in the GNP equation are jointly equal to zero can be rejected at better than the 1 percent level. In other words, the *Index of Consumer Sentiment* helps to predict GNP movements.

Table II also reports the sum of each block of coefficients and the standard error of the sum. The sum of consumer sentiment coefficients is positive in all GNP equations and significantly different

9. On causality in econometrics see Geweke, Meese and Dent [1983] and the September/October 1988 special issue of the *Journal of Econometrics*.

10. All reported results are drawn from equations that also include a constant term.

11. Here and below, the *Index of Consumer Sentiment* is scaled by a factor of 1,000 for clarity of presentation.

TABLE II
Vector Autoregressions with GNP, and the Index of Consumer Sentiment

	1 lag Dependent variable		2 lags Dependent variable		3 lags Dependent variable		4 lags Dependent variable	
	GNP	ICS	GNP	ICS	GNP	ICS	GNP	ICS
GNP								
F: p-value	0.118	0.206	0.563	0.314	0.761	0.086	0.808	0.063
Sum of coefficients	0.148	−0.066	0.063	−0.031	−0.085	−0.165+	−0.066	−0.263*
Standard error	(0.094)	(0.052)	(0.124)	(0.072)	(0.158)	(0.093)	(0.198)	(0.108)
ICS								
F: p-value	0.002	0.000	0.003	0.000	0.000	0.000	0.001	0.000
Sum of coefficients	0.265**	0.946**	0.272**	0.936**	0.314**	0.979**	0.327**	0.995**
Standard error	(0.082)	(0.045)	(0.087)	(0.050)	(0.096)	(0.056)	(0.107)	(0.058)
$\overline{R}^2$	0.162	0.825	0.157	0.831	0.215	0.838	0.202	0.845

This table reports information from vector autoregressions $Z_t = \Phi(L)Z_t + \varepsilon_t$ where Z_t contains GNP, and consumer sentiment. The 1-, 2-, 3-, and 4-lags models contain 121, 115, 110, and 109 observations, respectively. Each column is a regression. The dependent variable is indicated at the top of each column. In the main entries, the first number is the p-value associated with the F-statistic for the hypothesis that the block of coefficients are jointly equal to zero. Below this is the sum of the block of coefficients and its standard error. Significance levels on the sum of coefficients are indicated as follows: "**" is significant rejection at 1 percent, "*" is significant rejection at 5 percent, and "+" is significant rejection at 10 percent.

from zero at better than the 1 percent level. There is a subtle difference between this information and that provided by the F-statistic. The F-statistic looks at the role of the consumer sentiment in explaining quarterly GNP fluctuations while the sum of coefficients represents the "long-run" effect of consumer sentiment on the level of GNP.

The sum and block of GNP coefficients in the GNP equations are never significantly different from zero. The GNP coefficients in the *Index of Consumer Sentiment* equations are always negative and marginally significant in the three- and four-lags models. This curious effect recurs throughout the estimates below. Finally, the *Index of Consumer Sentiment* does an excellent job predicting its own movements.

A variant of the Granger test, described by Geweke, Meese, and Dent [1983], evaluates whether a possible forcing variable x decreases the forecast variance of variable y. Let the residual variance of y in an autoregression be σ^2. Let the residual variance when x is included as an explanatory variable be φ^2. Let the number of observations be N and the number of lags k. Geweke, Meese, and Dent show that

$$\eta(k) = N\left((\sigma^2 - \varphi^2)/\sigma^2\right) \sim \chi_k^2$$

and

$$\mu(k) = N\left((\sigma^2 - \varphi^2)/\varphi^2\right) \sim \chi_k^2$$

asymptotically under the null hypothesis of no Granger causality. They argue that

TABLE III
Forecast Error Tests for GNP and the Index of Consumer Sentiment

Lags (k)	σ^2	φ^2	$\eta(k)$	Observations
1	0.0000895	0.0000829	8.89	121
2	0.0000855	0.0000785	9.42	115
3	0.0000871	0.0000752	15.01	110
4	0.0000886	0.0000769	14.44	109

This table reports evidence on whether consumer sentiment reduces the forecast variance of GNP. Here σ^2 is the residual variance of GNP in an autoregression, φ^2 is the residual variance when consumer sentiment is included as an explanatory variable, and $\eta = N((\sigma^2 - \varphi^2)/\sigma^2)$. The null hypothesis is that expectations do not help to forecast GNP. The η statistics all indicate significant rejection at better than the 1 percent level.

this type of test has better small-sample properties than the Granger test above. The results for η are presented in Table III. In all four models the consumer sentiment index decreases the forecast variance of GNP at better than the 1 percent level of significance. Because $\mu(k) > \eta(k)$ in small samples, this result is even stronger using the statistic μ.

Output, Consumer Sentiment, and Leading Indicators

Tables II and III show that the *Index of Consumer Sentiment* helps to predict GNP, and thus by definition that GNP is Granger-caused by consumer sentiment. However, bivariate estimates do not make a compelling case for true causality between consumer sentiment and GNP because they fail to address the obvious possibility that the correlation is driven by a third variable acting on both consumer sentiment and GNP. Suppose that in reality GNP is caused entirely by a production shifter, say the price of oil, but with a lag of six months. How do we know that consumer sentiment is not simply a forecast of GNP based on observation of the price of oil? Because there are an infinite number of potential third variables, in general it is impossible to prove that a correlation is not caused by some omitted third variable. However, it is possible to evaluate if it is caused by specific variables. Our approach is to introduce into the vector autoregressions a succession of plausible variables that consumers might be using to forecast output and examine if they can eliminate the consumer sentiment—GNP correlation. To the extent that these third variables are unable to account for the correlation, the more confident we can be that the consumer sentiment—GNP relation is not simply a forecast.

The next step is then to estimate vector autoregressions including control variables that are correlated with economic fundamentals, that is,

$$\begin{bmatrix} Y_t \\ A_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) & c(L) \\ d(L) & e(L) & f(L) \\ g(L) & h(L) & i(L) \end{bmatrix} \begin{bmatrix} Y_t \\ A_t \\ S_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{Yt} \\ \varepsilon_{At} \\ \varepsilon_{St} \end{bmatrix},$$

where A_t represents the control variables. Series that are exogenous and have a good ability to forecast GNP are desirable. Inclusion of a large number of series in a vector autoregression consumes degrees of freedom and complicates statistical inference, so we choose a single series incorporating the effects of a number of different predictors, the *Index of Leading Indicators*, which is explicitly constructed to forecast GNP movements. The *Leading Indicators* is a composite index that includes a number of series that might be considered forcing variables (money supply, sensitive materials prices). It also contains endogenous variables (inventories, unemployment claims). Because the *Leading Indicators* contains both exogenous and endogenous variables, one expects the relationship between the *Leading Indicators* and *Index of Consumer Sentiment* to exhibit bidirectional causality if sentiment is causal. That is, the hypothesis that consumer sentiment causes output is not inconsistent with the possibility that it also reflects current information.

The results are reported in Table IV. The table is formatted in the same way as Table II. There are three pieces of evidence on the consumer sentiment—GNP relation. First, p-values are presented corresponding to the F-statistic on the block of consumer sentiment coefficients in the GNP equations, that is, the test of $c(L) \equiv 0$. The hypothesis that the consumer sentiment coefficients are jointly zero can be rejected at better than the 5 percent level in all models. The sum of consumer sentiment coefficients is positive and significantly different from zero at the 1 percent level in all models. The third relevant statistic, reported at the bottom of the table, tests for the block exogeneity of GNP and the *Leading Indicators* with respect to the *Index of Consumer Sentiment*. A χ^2 is constructed by scaling the difference between the log determinant of the residual covariance matrix of the estimated system and that of a bivariate GNP—*Leading Indicators* system. This statistic allows for the possibility that the *Index of Consumer Sentiment* affects GNP *indirectly* through the *Leading Indicators*. The hypothesis that GNP and the *Leading Indicators* are block exogenous can be strongly rejected in the one-, two-, and four-lags models. The χ^2 from the three-lags model does not attain statistical significance at conventional levels, but it is the sole exception to otherwise uniform evidence that the *Index of Consumer Sentiment* predicts GNP, and as we discuss below, the three-lags model can be statistically rejected in favor of the four-lags model.

Turning to other coefficients, it can be seen that GNP does a poor job predicting its own innovations, as above. Not surprisingly, the *Leading Indicators* predicts GNP. In the *Leading Indicators* equations, GNP appears to have a negative effect while the *Index of Consumer Sentiment* has a positive effect after three or four quarters. The consumer sentiment equation indicates that consumer sentiment has a substantial autoregressive component. The *Leading Indicators* has a positive effect, suggesting that consumer sentiment is in part a forecast based on fundamentals. However, a comparison with Table II reveals that inclusion of the *Leading Indicators* does not add much to the consumer sentiment $\bar{R}^2$. GNP has a negative effect, suggesting a kind of regression to the mean in expectations; when times have been good for a while people believe all good things must end and when times have been bad they believe there is nowhere to go but up.

Consumer sentiment appears to have a statistically significant effect on GNP. We would like to know if it has a *quantitatively* significant effect. Comparing the estimates in Table IV with those for bivariate vector autoregressions of GNP and the *Leading Indicators* (not included) shows that addition of consumer sentiment adds only 2 percent to 4 percent to the $\bar{R}^2$. The *Leading Indicators* and *Index of Consumer Sentiment* move together, so this crude estimate is probably a lower bound on the marginal contribution of sentiment to explaining GNP fluctuations.

A better way to assess the quantitative importance of consumer confidence in GNP fluctuations is to use the vector autoregression residuals to decompose the

TABLE IV
Vector Autoregressions with GNP, the Index of Leading Indicators, and the Index of Consumer Sentiment

	1 lag			2 lags			3 lags			4 lags		
	Dependent variable			*Dependent variable*			*Dependent variable*			*Dependent variable*		
	GNP	ILI	ICS	GNP	ILI	ICS	GNP	ILI	ICS	GNP	ILI	ICS
GNP												
F: p-value	0.149	0.006	0.002	0.069	0.046	0.011	0.233	0.175	0.006	0.228	0.126	0.009
Sum of coefficients	−0.138	−0.592**	−0.182**	−0.257+	−0.913**	−0.111	−0.327+	−1.161*	−0.259*	−0.397	−1.126*	−0.316*
Standard error	(0.095)	(0.213)	(0.056)	(0.137)	(0.329)	(0.086)	(0.195)	(0.463)	(0.116)	(0.242)	(0.535)	(0.133)
ILI												
F: p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sum of coefficients	0.249**	0.656**	0.101**	0.298**	0.659**	0.101**	0.290**	0.700**	0.110*	0.318**	0.369+	0.078
Standard error	(0.041)	(0.092)	(0.024)	(0.054)	(0.130)	(0.033)	(0.075)	(0.179)	(0.044)	(0.092)	(0.205)	(0.051)
ICS												
F: p-value	0.004	0.396	0.000	0.010	0.387	0.000	0.042	0.544	0.000	0.045	0.052	0.000
Sum of coefficients	0.214**	0.138	0.925**	0.236**	0.246	0.920**	0.260**	0.299	0.955**	0.280**	0.474*	0.980**
Standard error	(0.073)	(0.162)	(0.043)	(0.075)	(0.181)	(0.045)	(0.088)	(0.208)	(0.052)	(0.098)	(0.218)	(0.054)
$\bar{R}^2$	0.355	0.313	0.846	0.363	0.279	0.861	0.347	0.258	0.864	0.339	0.336	0.869
χ^2: p-value (Block exogeneity)		0.003			0.003			0.273			0.022	

This table reports information from vector autoregressions $Z_t = \Phi(L)Z_t + \varepsilon_t$ where Z_t contains GNP, consumer sentiment, and the leading indicators. The 1-, 2-, 3-, and 4-lags models contain 121, 115, 110, and 109 observations, respectively. Each column is a regression. The dependent variable is indicated at the top of each column. In the main entries, the first number is the p-value associated with the F-statistic for the hypothesis that the block of coefficients are jointly equal to zero. Below this is the sum of the block of coefficients and its standard error. Significance levels on the sum of coefficients are indicated as follows: "**" is significant rejection at 1 percent, "*" is significant rejection at 5 percent, and "+" is significant rejection at 10 percent. The χ^2 statistic is for the hypothesis that GNP and the leading indicators are block exogenous with respect to consumer sentiment.

forecast variance of GNP into contributions by each of the variables. The technique is briefly outlined below; a comprehensive discussion can be found in Sims [1980a] (see also Sims [1980b] and Litterman and Weiss [1985].) As a caveat to this exercise we should repeat Schiller's [1987] observation that variables with little ability to explain long-run variance nevertheless may be highly important in particular situations.

Let u be a three-element vector of forecast errors for a trivariate vector autoregression and Σ the covariance matrix associated with the u process. A corresponding orthonormal vector v and a lower triangular matrix G can be found such that $GG' = \Sigma$ and $Gv = u$. Because v is orthonormal and G is lower triangular the variance of the forecast error of each of the endogenous variables can be decomposed into percentage contributions by innovations in each of the three equations. This procedure can be used to decompose variances of any arbitrary length forecast.

The percentage contributions are sensitive to the decomposition method, specifically to the ordering of variables in the equation. For example, by construction the first variable listed contributes 100 percent of its own one-period-ahead forecast variance. Because G is lower triangular it is conventional to place forcing variables early in the order and endogenous variables later. For this reason the hypothesized driving variables, the *Index of Consumer Sentiment* and the *Leading Indicators*, are placed first in the first two decompositions. Consumer sentiment is placed last, somewhat atheoretically, in the third decomposition to describe the worst case.

Decompositions of GNP and the *Index of Consumer Sentiment* from one to eight quarters ahead are presented in Table V.[12] The estimates are for the four-lags model in Table IV. Here and below, when choosing which model to decompose we were guided by the following selection rule. First we computed the distance between the residual covariance matrixes of the three- and four-lags models. If they were significantly different, we chose the four-lags model. If not, we compared the two- and three-lags models. If they were significantly different, we chose the three-lags model, and so on. The variables are listed in the order they appear in the u vector.

In the top GNP decomposition, the variance of the *Index of Consumer Sentiment* explains 14.4 percent of the one-quarter-ahead forecast variance of GNP, rising to 25.8 percent eight quarters ahead. In the bottom decomposition, which is the worst case for finding an effect of consumer sentiment, consumer sentiment explains 12.9 percent of the eight-quarter-ahead variance. Recall that the zero percent estimate of the one-quarter-ahead contribution in the last equation is entirely an artifact of the decomposition technique. Rough bounds on the mean effect of consumer sentiment on GNP then would be between 13 percent and 26 percent. It seems reasonable to conclude that movements in consumer confidence are quantitatively important in explaining GNP fluctuations.

The consumer sentiment decompositions provide evidence on the exogeneity of consumer sentiment. In the top panel, the *Index of Consumer Sentiment* explains 83.6 percent of its own innovation variance eight quarters ahead. In the worst case for the *Index*, the bottom panel, consumer sentiment still explains 72.1 percent of its eight-quarter-ahead innovation variance. Although consumer sentiment appears to be affected by GNP and leading indicator movements, neither of these variables can explain the majority of sentiment variation. This suggests that the *Index of Consumer Sentiment* is to a large degree exogenous with respect to these

12. RATS version 4.02 was used to perform the variance decompositions. The routine ignores the sampling error component of the process and considers only the equation error.

TABLE V
Decomposition of the Innovation Variance of GNP and the Index of Consumer Sentiment from Vector Autoregressions with GNP, Consumer Sentiment, and the Leading Indicators

	GNP decomposition			*ICS decomposition*		
Quarter	ICS	ILI	GNP	ICS	ILI	GNP
1	14.4	10.8	74.8	100.0	0.0	0.0
4	21.1	21.8	57.1	90.1	2.8	7.1
8	25.8	21.7	52.5	83.6	3.2	13.2
	GNP decomposition			*ICS decomposition*		
Quarter	ILI	ICS	GNP	ILI	ICS	GNP
1	22.5	2.7	74.8	25.2	74.8	0.0
4	39.0	3.9	57.1	32.5	60.4	7.1
8	35.2	12.3	52.5	22.4	64.5	13.2
	GNP decomposition			*ICS decomposition*		
Quarter	GNP	ILI	ICS	GNP	ILI	ICS
1	100.0	0.0	0.0	14.4	13.3	72.2
4	71.5	25.8	2.7	5.9	29.7	64.4
8	62.9	24.3	12.9	4.1	23.8	72.1

This table reports the percentage of GNP and consumer sentiment innovation variance that can be attributed to each variable in the four-lags vector autoregression. For each model, the decomposition is computed for three different orderings of the variables. The quarter (ahead) of the innovation variance being decomposed is indicated in the leftmost column. The remaining entries indicate the percentage of innovation variance attributable to the variable at the top of the column.

variables, and tends to undermine the notion that sentiment is nothing more than a forecast of output.

Robustness

The results in the previous section support the hypothesis that consumer sentiment contributes to fluctuations in aggregate output. However, like the test of Oh and Waldman [1990], the estimates can be interpreted in more than one way. In particular, there remains the possibility that we have omitted an economic fundamental that is forcing both the consumer sentiment and GNP series. This section explores a number of different specifications and variables in pursuit of such a fundamental.

First, we estimate a set of vector autoregressions with GNP, consumer sentiment, and three "textbook" control variables: money supply, government spending, and sensitive materials prices. We call these textbook controls because they are staples of almost every macroeconomics textbook. If one were to ask economists to name exogenous factors that cause GNP fluctuations, these variables would probably appear on most lists.

Table VI reports evidence from these vector autoregressions. To conserve space,

TABLE VI
Vector Autoregressions and Variance Decompositions of GNP, the Index of Consumer Sentiment, and Textbook Controls

	1 lag		2 lags		3 lags		4 lags	
	GNP	ICS	GNP	ICS	GNP	ICS	GNP	ICS
GNP								
F: p-value	0.722	0.194	0.792	0.214	0.751	0.163	0.870	0.299
Sum of coefficients	0.035	–0.070	–0.068	0.051	–0.239	–0.095	–0.205	–0.169
Standard error	(0.099)	(0.054)	(0.139)	(0.077)	(0.193)	(0.107)	(0.254)	(0.132)
ICS								
F: p-value	0.055	0.000	0.140	0.000	0.109	0.000	0.158	0.000
Sum of coefficients	0.161+	0.873**	0.194*	0.850**	0.267*	0.876**	0.236+	0.888**
Standard error	(0.083)	(0.045)	(0.090)	(0.051)	(0.110)	(0.062)	(0.135)	(0.070)
$\bar{R}^2$	0.266	0.850	0.283	0.871	0.285	0.874	0.271	0.871
χ^2: p-value	0.000		0.000		0.000		0.015	

	GNP decomposition			*ICS decomposition*		
Quarter	ICS	X	GNP	ICS	X	GNP
1	19.3	7.1	73.6	100.0	0.0	0.0
4	23.7	18.7	57.6	86.1	11.4	2.5
8	21.7	30.8	47.5	65.1	28.1	6.8
Quarter	X	ICS	GNP	X	ICS	GNP
1	9.5	16.9	73.6	6.6	93.4	0.0
4	23.8	18.6	57.6	24.0	73.5	2.5
8	35.0	17.5	47.5	37.1	56.1	6.8

This table reports information from vector autoregressions $Z_t = \Phi(L)Z_t + \varepsilon_t$ where Z_t contains GNP, consumer sentiment, and control variables. The controls are government spending, money supply M2, and sensitive materials prices. Each column in the top panel corresponds to a regression. Only the GNP and consumer sentiment regressions are reported, and only the GNP and consumer sentiment coefficients. The 1-, 2-, 3-, and 4-lags models contain 121, 115, 110, and 109 observations, respectively. The dependent variable is indicated at the top of each column. In the main entries the first number is the p-value associated with the F-statistic for the hypothesis that the block of coefficients are jointly equal to zero. Below this is the sum of the block of coefficients and its standard error. The χ^2 statistic is for the hypothesis that GNP and the controls are block exogenous with respect to consumer sentiment. Significance levels on the sum of coefficients are indicated as follows: "**" is significant rejection at the 1 percent level; "*" is significant at the 5 percent level; "+" is significant at 10 percent. The bottom panel reports the percentage of GNP and consumer sentiment innovation variance that can be attributed to each variable in the 4-lags vector autoregression. The sum of variance explained by the controls is reported under the heading "X."

only the GNP and consumer sentiment equations are reported, and only the coefficients on GNP and *Index of Consumer Sentiment*. The consumer sentiment coefficients in the GNP equation are of primary interest. The F-statistic on the block of consumer sentiment coefficients is significant at the 10 percent level in the one-lag model and just beyond conventional significance levels in the two-, three-, and four-lags models. Some fall in precision of estimates is to be expected, especially in the model with four lags, because of the loss of degrees of freedom. The sum of consumer sentiment coefficients is positive and statistically significant in all models. The magnitude of the sum in the one-lag and two-lags models is quite a bit lower than in Table IV, but about the same for the three- and four-lags model, indicating that the controls rob consumer sentiment of some power in the first two quarters.

The most important statistic is the χ^2 that tests for the block exogeneity of GNP and the controls with respect to the *Index of Consumer Sentiment*. Recall that this indicates whether consumer sentiment has any effect on the system, direct or indirect. Block exogeneity can be rejected at approximately the 1 percent level in all models. Taken together, the estimates suggest that the textbook set of controls cannot account for the consumer sentiment—GNP relation.

The bottom panel of Table VI reports variance decompositions. To maintain comparability with other tables we choose to present the decomposition of the four-lags model, although in this case our selection criteria suggested that the two- and three-lags models would be fine as well. In any case, the two- and three-lags decompositions are substantially the same. As in Table V we decompose GNP and the *Index of Consumer Sentiment*. The sum of the percentage contribution of controls is indicated in the column headed "X." Again, to conserve space, only two orderings are reported. In the GNP equation, when consumer sentiment is ordered before the controls, consumer sentiment can account for over 20 percent of the innovation variance of output, with its explanatory power peaking three and four quarters ahead at 23.7 percent. When consumer sentiment is listed after the controls, it experiences only a minor decline—four quarters ahead it still explains 18.6 percent of GNP innovation variance. In the consumer sentiment equation, the *Index* accounts for no less than 65 percent of its own innovation variance when listed first, and no less than 56 percent when listed after all the controls. The decompositions from the textbook system do not differ in substance from the *Leading Indicators* system: consumer sentiment has a quantitatively large effect on GNP and appears to have an important exogenous component, suggesting it is not simply a forecast of GNP.

The next set of vector autoregressions include as controls both the *Leading Indicators* and default risk, the spread between the interest rate on corporate debt and treasury bills. The default risk variable has attracted recent attention for its ability to predict GNP movements (Stock and Watson [1989]). It is not clear whether this variable reflects an economic fundamental or is itself a non-fundamental, perhaps the "animal spirits" of investors.

Table VII reports vector autoregressions with GNP, consumer sentiment, the *Leading Indicators*, and default risk, indicated "DR" in the table. For space reasons, only the four-lags model is reported. Our test for number of lags suggests that models with fewer than four lags can be rejected at about the 1 percent level. The key statistics, as usual, concern consumer sentiment in the GNP equation. First, the hypothesis that the block of consumer sentiment coefficients are jointly zero can be rejected at almost the 5 percent level. Second, the sum of coefficients is positive and statistically different from zero at better

TABLE VII
Vector Autoregressions with GNP, the Index of Leading Indicators, the Index of Consumer Sentiment, and Default Risk

	Dependent variable GNP	ILI	ICS	DR
GNP				
F: *p*-value	0.177	0.038	0.020	0.516
Sum of coefficients	−0.441+	−1.316**	−0.296*	0.028
Standard error	(0.242)	(0.483)	(0.130)	(0.097)
ILI				
F: *p*-value	0.118	0.313	0.016	0.183
Sum of coefficients	0.205+	0.061	0.050	0.066
Standard error	(0.111)	(0.221)	(0.060)	(0.044)
ICS				
F: *p*-value	0.063	0.369	0.000	0.001
Sum of coefficients	0.230*	0.272	0.971	0.039
Standard error	(0.099)	(0.198)	(0.054)	(0.040)
DR				
F: *p*-value	0.060	0.000	0.010	0.000
Sum of coefficients	−0.661+	−2.603**	−0.117	0.754**
Standard error	(0.396)	(0.791)	(0.213)	(0.159)
$\bar{R}^2$	0.374	0.489	0.881	0.578
χ^2: *p*-value (Block exogeneity)		0.011		

This table reports information from a vector autoregression $Z_t = \Phi(L)Z_t + \varepsilon_t$ where Z_t contains GNP, consumer sentiment (ICS), leading indicators (ILI), and default risk (DR). The model has 4 lags and 109 observations. The dependent variable is indicated at the top of each column. In the main entries, the *F*-statistic is for the hypothesis that the block of coefficients are jointly equal to zero. Beneath the *F* is the sum of coefficients and its standard error. The χ^2 statistic is for the hypothesis that GNP, the leading indicators, and default risk are block exogenous with respect to consumer sentiment. Significance levels on the sum of coefficients are indicated as follows: "**" is significant rejection at 1 percent, "*" is significant rejection at 5 percent, and "+" is significant rejection at 10 percent.

than the 5 percent level. Addition of the default risk variable leads to a modest decrease in the magnitude of the sum. The χ^2 statistic indicates that block exogeneity of GNP, the *Leading Indicators*, and default risk with respect to consumer sentiment can be rejected at roughly the 1 percent level. The evidence continues to point to a role for consumer confidence in economic fluctuations.[13]

13. The χ^2 statistic and sum of coefficients are also statistically significant in the unreported one-, two-, and three-lags models, while the *F*-statistics lie just outside statistical significance.

For this model, we report all of the equations and coefficients because the default risk variable is of some interest in itself. This variable has a significant effect on the movement of all variables in the system according to the *F*-statistic. The sum of default risk coefficients is significant in the GNP, *Leading Indicators*, and default risk equations. Thus, default risk matters.

This is interesting because default risk is not obviously an economic fundamental. It may be simply a good mirror of underlying fundamentals, but the failure of the *Leading Indicators* and GNP to

TABLE VIII
Decomposition of the Innovation Variance of GNP and the Index of Consumer Sentiment from Vector Autoregressions with GNP, Consumer Sentiment, the Leading Indicators, and Default Risk

	GNP decomposition				*ICS decomposition*			
Quarter	ICS	ILI	DR	GNP	ICS	ILI	DR	GNP
1	11.7	4.9	0.0	83.3	100.0	0.0	0.0	0.0
4	14.5	11.3	18.1	56.1	86.3	1.5	7.3	4.9
8	26.9	10.8	15.0	47.2	85.9	1.3	4.5	8.3
	GNP decomposition				*ICS decomposition*			
Quarter	ILI	ICS	DR	GNP	ILI	ICS	DR	GNP
1	11.8	4.9	0.0	83.3	16.7	83.3	0.0	0.0
4	19.4	6.3	18.1	56.1	21.0	66.9	7.3	4.9
8	16.5	21.3	15.0	47.2	14.9	72.3	4.5	8.3
	GNP decomposition				*ICS decomposition*			
Quarter	DR	ILI	GNP	ICS	DR	ILI	GNP	ICS
1	2.0	9.7	88.2	0.0	2.1	14.7	4.7	78.5
4	27.4	9.7	58.4	4.5	15.1	12.0	2.0	70.9
8	22.3	8.8	47.2	21.8	9.5	9.1	1.8	79.6

This table reports the percentage of GNP and consumer sentiment innovation variance that can be attributed to each variable in the 4-lags vector autoregression. The decomposition is computed for three different orderings of the variables. The quarter (ahead) of the innovation variance being decomposed is indicated in the leftmost column. The remaining entries indicate the percentage of innovation variance attributable to the variable at the top of the column.

achieve statistical significance in the default risk equation does not square well with this interpretation. The results were no better in the models with fewer lags. In fact, the only variable identified as driving default risk in this system is consumer sentiment. Taken together, the evidence is not inconsistent with the idea that default risk represents investor sentiment, or captures some other aspect of consumer confidence.

Table VIII reports the GNP and consumer sentiment variance decompositions for the model in Table VII. The evidence follows the pattern established above. Consumer sentiment accounts for a healthy fraction of the innovation variance of GNP, up to 21.8 percent eight quarters ahead in the worst case decomposition ordering. Default risk serves primarily to rob the *Leading Indicators* of its explanatory power. Furthermore, consumer sentiment continues to appear largely exogenous. Even in the worst case decomposition, consumer sentiment accounts for no less than 70 percent of its own innovation variance. If anything, inclusion of default risk in the vector autoregressions suggests an even more important role for consumer confidence as a causal factor.

In the final set of vector autoregressions, both default risk and the individual components of the *Leading Indicators* are included as controls.[14] The benefit of this is that it avoids the ad hoc parameter restrictions implied by construction of the *Leading Indicators*. The cost is that the large number of component series rapidly consume the available degrees of freedom.

The GNP and consumer sentiment coefficients for the GNP and consumer sentiment equations are reported in the upper panel of Table IX. Variance decompositions are reported in the bottom panel. Beginning as usual with the *F*-statistic on the block of consumer sentiment coefficients in the GNP equation, the hypothesis that consumer sentiment does not Granger-cause GNP can be rejected at about the 5 percent level in the one-lag and three-lags models, but does not quite achieve statistical significance in the two- and four-lags models. Here again, this test loses a fair amount of power due to the large number of control variables, especially in the model with four lags. The sum of the consumer sentiment coefficients is positive and significantly different from zero in all models. Most important, the block exogeneity of GNP and the controls with respect to consumer sentiment can be rejected at better than the 1 percent level in all four models.

The bottom panel of Table IX reports variance decompositions for the four-lags model. Here again the models with fewer than four lags can be rejected when comparing the log determinants of the residual covariance matrixes. Two decomposition orderings are reported, one that lists consumer sentiment before the controls (indicated as "X") and one that lists it after the controls; GNP is ordered last in both cases. In the first GNP decomposition, consumer sentiment explains 21.6 percent of the one-quarter-ahead variance; the effect falls gradually to 16.3 percent eight quarters ahead. In the second GNP decomposition, the estimated consumer sentiment effect is lower, peaking at 8.8 percent five quarters ahead. These effects are smaller than in previous tables, but not trivial. In the consumer sentiment decompositions, the evidence continues to support the idea that the *Index of Consumer Sentiment* has a substantial exogenous component, that is, consumer sentiment is more than a forecast based on available information. In the first decomposition, consumer sentiment explains no less than 52 percent of its own innovation variance, while in the second no less than 31 percent.

The results in this section can be simply summarized. Even after controlling for a number of good predictors of GNP, consumer sentiment continues to add predictive power to the model. These results cannot rule out the possibility that consumer sentiment is nothing more than a forecast of output. However, under this interpretation it appears that consumers are basing their forecasts on information that is unavailable to professional econometricians who specialize in forecasting GNP movements. The alternative interpretation that consumer confidence causes output movements seems no less plausible.

V. CONCLUSION

This paper explores the possibility that the economy's total output occasionally varies not in response to a shift in fundamentals but in response to a shift in consumer sentiment. Specifically, the paper asks whether and to what extent exogenous declines in consumer confidence cause recessions and conversely whether and to what extent bullish consumers drive economic growth.

14. To be precise, we use the components of the current *Leading Indicators* less the *Index of Consumer Expectations*. The current *Leading Indicators* differs from the version listed in Table I in that it includes consumer expectations and unfilled orders and excludes the change in manufacturing and trade inventories.

TABLE IX
Vector Autoregressions and Variance Decompositions of GNP, the Index of Consumer Sentiment, Default Risk, and Leading Indicators Components

	1 lag		2 lags		3 lags		4 lags	
	GNP	ICS	GNP	ICS	GNP	ICS	GNP	ICS
GNP								
F: p-value	0.014	0.048	0.031	0.123	0.036	0.178	0.324	0.571
Sum of coefficients	–0.308*	–0.137*	–0.482*	–0.024	–0.820*	–0.292+	–0.855+	–0.263
Standard error	(0.123)	(0.069)	(0.230)	(0.123)	(0.328)	(0.168)	(0.498)	(0.250)
ICS								
F: p-value	0.053	0.000	0.149	0.000	0.047	0.000	0.254	0.000
Sum of coefficients	0.237+	0.781**	0.278+	0.822**	0.462**	0.942**	0.545*	0.855**
Standard error	(0.121)	(0.067)	(0.160)	(0.089)	(0.205)	(0.110)	(0.272)	(0.137)
$\overline{R}^2$	0.336	0.867	0.330	0.874	0.369	0.889	0.315	0.887
χ^2: p-value	0.001		0.000		0.000		0.000	

	GNP decomposition			*ICS decomposition*		
Quarter	ICS	X	GNP	ICS	X	GNP
1	21.6	31.1	47.3	100.0	0.0	0.0
4	16.3	54.9	28.8	64.8	34.6	0.6
8	16.3	61.8	21.9	52.9	46.6	0.5
Quarter	X	ICS	GNP	X	ICS	GNP
1	48.0	4.6	47.4	39.7	60.3	0.0
4	64.1	7.1	28.8	59.6	39.8	0.6
8	70.3	7.8	21.9	67.8	31.7	0.5

This table reports information from vector autoregressions $Z_t = \Phi(L)Z_t + \varepsilon_t$ where Z_t contains GNP, consumer sentiment, and control variables. The controls are default risk and the components of the current *Leading Indicators*, except for the *Index of Consumer Expectations*. The 1-, 2-, 3-, and 4-lags models contain 114, 112, 110, and 109 observations, respectively. Each column in the top panel corresponds to a regression. Only the GNP and consumer sentiment regressions are reported, and only the GNP and consumer sentiment coefficients. The dependent variable is indicated at the top of each column. In the main entries the first number is the p-value associated with the F-statistic for the hypothesis that the block of coefficients are jointly equal to zero. Below this is the sum of the block of coefficients and its the standard error. The χ^2 statistic is for the hypothesis that GNP and the controls are block exogenous with respect to consumer sentiment. Significance levels for the sum of coefficients are indicated as follows: "**" is significant rejection at the 1 percent level; "*" is significant at the 5 percent level; "+" is significant at 10 percent. The bottom panel reports the percentage of GNP and consumer sentiment innovation variance that can be attributed to each variable in the vector autoregressions. The sum of variance explained by the controls is reported under the heading "X."

There are two inspirations for this research. The first is the fact that something called "consumer confidence" plays an important role in popular explanations of the business cycle and in the public statements of business and political leaders. There is a long tradition of confidence explanations for recessions. Keynes emphasized the "mass psychology of the market" and "animal spirits" among investors, concluding that "[i]n estimating the prospects of investment, we must have regard, therefore, to the nerves and hysteria and even the digestions and reactions to the weather of those upon whose spontaneous activity it largely depends." During the Great Depression President Franklin D. Roosevelt and General Motors Chairman Alfred P. Sloan, Jr., who disagreed in many ways on the role of government in helping the economy, seemed to be of the same mind on the nature of the problem and the primary impediment to recovery. Roosevelt's famous inaugural speech contained the statement, "let me assert my firm belief that the only thing we have to fear is fear itself" while Sloan commented, "the main obstacle to the general revival of American enterprise is the fear the foundation of the economy is in jeopardy."[15]

Our empirical estimates for the United States, 1953–1988, uniformly reject the hypothesis that consumer sentiment does not cause GNP (in the Granger sense). This gives support to the idea that exogenous changes in consumer sentiment have real effects on output. According to our central vector autoregressions, between 13 percent and 26 percent of the variance of GNP innovations is the result of waves of consumer sentiment, which suggests a nontrivial role for expectations. Thus, the main contribution of the paper is a demonstration that consumer confidence is an important independent factor in economic fluctuations, a fact that did not previously rest on solid empirical ground (Leeper [1991]).

The second purpose of the paper is to provide some evidence on the rich collection of macroeconomic models with strategic complementarities that have been developed in recent years. All these multiple-equilibria models have in common that expectations are self-fulfilling—because agents must expect to be in a particular equilibrium before the economy can move to it, their expectations in a sense cause the movement to the equilibrium. An implication of these models, then, is that after controlling for movements in economic fundamentals, changes in consumer sentiment lead to changes in GNP. Our estimates appear to provide support for this key implication.

However, we caution that our evidence does not prove that expectations cause output fluctuations. There remains the possibility that a missing third variable is forcing both series. Even so, the results have power because we could have found that expectations were completely insignificant. If that had been the case, it would have been safe to conclude that multiple equilibria are not empirically important. Like Oh and Waldman [1990], then, our results are open to an alternative interpretation. The two studies taken together constitute an initial case in favor of the class of macroeconomic models that view some recessions as coordination failures, but certainly additional research is called for.

15. Quotations in the paragraph are from Keynes [1964, 154, 161, 162], Roosevelt [1946, 13] and Cray [1980, 311], respectively.

REFERENCES

Ball, Laurence, and David Romer. "Sticky Prices as Coordination Failure." *American Economic Review*, June 1991, 539–52.

Bryant, John. "A Simple Keynes-Type Model." *The Quarterly Journal of Economics*, August 1983, 525–28.

______. "The Paradox of Thrift, Liquidity Preference, and Animal Spirits." *Econometrica*, September 1987, 1231–35.

Cass, David, and Karl Shell. "Do Sunspots Matter?" *Journal of Political Economy*, April 1983, 193–227.

Cray, Ed. *Chrome Colossus: General Motors and Its Times*. New York, N.Y.: McGraw-Hill Book Company, 1980.

Cooper, Russell, and Andrew John. "Coordinating Coordination Failures in Keynesian Models." *The Quarterly Journal of Economics*, August 1988, 441–63.

Curtin, Richard T. "Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers." *Public Opinion Quarterly*, Fall 1982, 340–52.

Dagsvik, John, and Boyan Jovanovic. "Was the Great Depression a Low-Level Equilibrium?" C. V. Starr Center for Applied Economics, R. R. 91–07, January 1991.

Friend, Irvin, and F. Gerard Adams. "The Predictive Ability of Consumer Attitudes, Stock Prices, and Non-Attitudinal Variables." *Journal of the American Statistical Association*, December 1964, 987–1005.

Geweke, John, Richard Meese, and Warren Dent. "Comparing Alternative Tests of Causality in Temporal Systems." *Journal of Econometrics*, 21, February 1983, 161–94.

Hertzberg, Marie P., and Barry A. Beckman. "Business Cycle Indicators: Revised Composite Indexes." *Business Conditions Digest*, January 1989, 97–102.

İmrohoroglu, Selahattin. "Testing for Sunspot Equilibria in the German Hyperinflation." *Journal of Economic Dynamics and Control*, 17, January/March 1993, 289–317.

Keynes, John Maynard. *The General Theory of Employment, Interest, and Money*. New York, N.Y.: Harcourt Brace Jovanovich, 1964.

Kiyotaki, Nobuhiro. "Multiple Expectational Equilibria Under Monopolistic Competition." *The Quarterly Journal of Economics*, November 1988, 695–713.

Leeper, Eric M. "Consumer Attitudes and Business Cycles." Manuscript, Federal Reserve Bank of Atlanta, August 1991.

Litterman, Robert B., and Lawrence Weiss. "Money, Real Interest Rates and Output: A Reinterpretation of Postwar U.S. Data." *Econometrica*, 53, January 1985, 129–56.

Milgrom, Paul, and John Roberts. "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica*, November 1990, 1255–77.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. "Industrialization and the Big Push." *Journal of Political Economy*, October 1989, 1003–26.

Oh, Seonghwan, and Michael Waldman. "The Macroeconomic Effects of False Announcements." Manuscript, *The Quarterly Journal of Economics*, November 1990, 1017–34.

______. "The Leading Indicators as a Source of Expectational Shocks." Manuscript, UCLA and Cornell, 1993.

Roberts, John. "An Equilibrium Model with Involuntary Unemployment at Flexible, Competitive Prices and Wages." *American Economic Review*, December 1987, 856–74.

Roosevelt, Franklin D. *Nothing to Fear: The Selected Addresses of Franklin D. Roosevelt 1932–1945*, edited by B. D. Zevin. New York, N.Y.: Houghton Mifflin Company, 1946.

Schiller, Robert J. "Ultimate Sources of Aggregate Variability." *American Economic Review*, May 1987, 77–92.

Shleifer, Andrei. "Implementation Cycles." *Journal of Political Economy*, December 1986, 1163–90.

Sims, Christopher A. "Macroeconomics and Reality." *Econometrica*, January 1980a, 1–48.

______. "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered." *American Economic Review*, May 1980b, 250–57.

Stock, James H., and Mark W. Watson. "New Indexes of Coincident and Leading Economic Indicators." *NBER Macroeconomics Annual*, edited by Olivier Jean Blanchard and Stanley Fischer. National Bureau of Economic Research, Cambridge, Mass., 1989, 351–94.

Weil, Philippe. "Increasing Returns and Animal Spirits." *American Economic Review*, September 1989, 889–94.

Woodford, Michael. "Three Questions about Sunspot Equilibria." *American Economic Review*, May 1987, 93–98.

Zarnowitz, Victor. *Orders, Production, and Investment—A Cyclical and Structural Analysis*, New York, N.Y.: NBER and Columbia University Press, 1973.