Tracking the Economy with the Purchasing Managers' Index

by Ethan S. Harris

In the last several years the purchasing managers' index has emerged as a key indicator of manufacturing activity. This "index" consists of five separate indexes measuring monthly changes in manufacturing output, employment, new orders, inventories, and vendor deliveries, together with a composite index that gives a weighted average of the other five. Financial markets are now quite sensitive to the index, and news reports on the economy regularly feature it. The index receives such close attention for several reasons: it is the first broad indicator released each month, it covers the cyclically sensitive manufacturing sector (Chart 1), and the data are easy to interpret and are virtually never revised.

Despite the index's popular appeal and market-moving power, some skepticism about the utility of this indicator is warranted. It is not constructed with the scientific sampling and statistical methods that underlie most official macroeconomic series (see appendix). A qualitative measure of activity, it reports whether business has increased or decreased but makes no assessment of the strength of the change. Most important, the index has not been rigorously tested: although there is ample evidence that the index tracks the general ups and downs of the economy, analysts have not demonstrated that the purchasing managers' data yield information on the economy beyond that already provided by other indicators.

This article analyzes the strengths and weaknesses of the index as a forecasting tool. It begins by explaining how the index is constructed. The next section presents the basic correlations between the five component indexes and the economic aggregates they are supposed to track. The remainder of the article investigates the predictive power of the purchasing managers' data: Do the indexes lag or lead economic activity? Do they foreshadow turning points in the business cycle? Can the indexes improve on the forecasts of simple economic models or on consensus forecasts?

Our results give mixed support for the purchasing managers' index. One shortcoming is the index's tendency to pick up activity in the weeks preceding the month it is supposed to measure. Another limitation is that none of the components explains more than half of the monthly variation in the corresponding official statistics. Furthermore, the index is not a reliable leading indicator: it sends too many false signals and its lead time is too erratic to be of use in anticipating cyclical swings. Nevertheless, the index does add significantly to the explanatory power of simple econometric models and consensus forecasts. And it could be even more useful to forecasters if the sampling and statistical methodology were improved. Thus, although the index has some important limitations, with careful application it can be useful in forecasting economic activity.

Description

About the middle of each month the National Association of Purchasing Managers (NAPM) surveys roughly 300 association members representing twenty-one manufacturing industries in all fifty states. The survey asks each purchasing manager how the current level of five key economic indicators—production, new orders, employment, inventories, and vendor delivery time—

compares with the previous month's level.[1] The responses are simply "higher," "lower," or "the same." The unweighted percentage of firms in each category is then tabulated and a diffusion index is constructed by summing the percentage of positive responses and one-half of those responding "the same."[2] A reading above 50 percent in a diffusion index means that more firms are expanding activity than contracting activity. Finally, these data are seasonally adjusted and combined into a single weighted composite index.

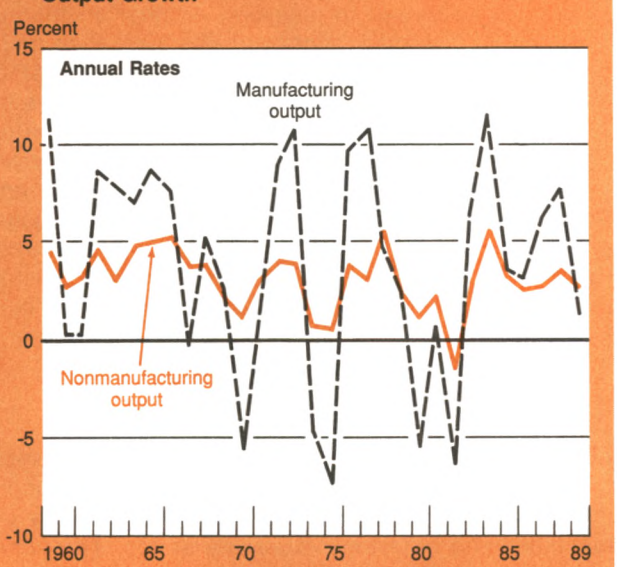
Although the survey has been published since 1931 (with an interruption for World War II), several of its more sophisticated features were only introduced in recent years. The data were originally published in raw, seasonally unadjusted form; in the early 1980s, with help from the Commerce Department, the association began publishing seasonally adjusted diffusion indexes. The sample size has also been increased to almost 300 from about 225. Since the summer of 1989, when financial markets became increasingly interested in the index, the survey has been released earlier and at the same time each month. It now usually "beats" the employment report by several days and thus captures maximum attention in the market.[3]

[1]The survey also includes questions on commodity prices and buying policy. In the last several years new export orders and imports have been added.

[2]The NAPM survey treats vendor delivery time somewhat differently. The responses for this indicator are "slower," "faster," and "no change." The diffusion index for vendor deliveries is the sum of the percentage reporting slower delivery time and half the percentage reporting no change.

Chart 1

Manufacturing and Nonmanufacturing Output Growth



Source: National Income and Product Accounts, Table 6.2.

The index as a measure of economic activity

The NAPM component indexes have counterparts in official data published by the federal government. Since the indexes are measures of the diffusion of the economic activity, they should have roughly a linear relationship to the *growth* in corresponding government data.[4] In other words, if a higher proportion of firms are reporting expanded activity, then we would expect higher growth in aggregate activity.

Table 1 presents evidence of how closely the NAPM data track the economy. The table shows the results of regressing the percent change in the official data on the corresponding component of the index. The first three columns present estimates using monthly data for the 1959-91 period. As the t-statistics (in parentheses) show, the NAPM coefficient is significant in all of the equations. The overall fit (R-square), however, is generally modest: the indexes explain less than half of the monthly variation in growth for all variables. The weakest results are for new orders and prices, two highly volatile series; the best results are for employment.

The fourth and fifth columns of Table 1 show the overall fit when quarterly and annual data are used. Although this time aggregation generally improves the fit, the NAPM data still leave a good portion of the variation in growth unexplained. The last two columns of Table 1 show the implied break-even point for the indexes. Theoretically, when aggregate growth is zero, a diffusion index should average out to about 50 percent, with equal numbers of firms reporting higher and lower activity. The regression estimates suggest that using 50 percent as the break-even point can be misleading. For example, in the regression of industrial production on the composite index, estimates for the 1980-91 period show a break-even rate of just 46.9 percent.[5]

[3]Interest in the index has also drawn attention to some of the regional purchasing managers' surveys. The Chicago index is closely watched, in part because it is released before the national index.

[4]Specifically, if (1) firms have identical but nonsynchronous cycles and (2) growth is evenly distributed among large and small firms along a rectangular distribution, then there will be an exact linear relationship between the proportion of firms expanding and the rate of growth of aggregate activity. Regression tests found no evidence of significant nonlinearities.

[5]Recent experience illustrates the danger of using 50 percent as the break-even point. From May 1989 to April 1990 the composite NAPM index dipped below 50 percent, averaging 47.6 percent. If

The index as a leading indicator

Tracing the general movements in economic activity is not a very rigorous test of an indicator. Much of the interest in the purchasing managers' index among business economists stems from its alleged ability to signal changes in economic trends. The tremendous attention the index now receives started in the summer of 1989 when the index, falling below 50 percent, appeared to presage a recession. Clearly the index's early release makes it a "timely indicator"; the more difficult question is whether it in fact anticipates activity in the months ahead. Does it lead activity or measure contemporaneous activity? And are business economists correct in assuming that it gives a reliable warning of recession?

The purchasing managers' index, like all diffusion indexes, has leading indicator qualities. Chart 2 shows the relationship between the composite index and the growth in manufacturing output over the business cycle. The index peaks when growth is highest, declines to 50 percent as growth levels off, and then falls below 50 percent as the economy slips into recession. Empirical work by Cox and Torda shows that the composite index "reached its cyclical peak about 11½ months before the onset of the seven postwar recessions" and that "the lead time of the composite index of leading economic indicators is similar, about 12 months."[6] Cox and Torda also find that the composite index generally leads cyclical recoveries.

Unfortunately, average lead time is a poor criterion for judging a leading indicator. To be useful, a leading indicator must predict turning points with a relatively regular lead of at least a few months. It must also give a relatively small number of false signals. Here we test the predictive power of two types of movement in the composite index: turning points in the index and periods when the index crosses various "break-even" or "threshold" points.

Neither NAPM signal reliably predicts business cycle turning points. As Chart 2 shows, the index often turns down long before a business cycle peak, reflecting the slowing of growth following the initial cyclical recovery. Even if we ignore this initial peak, the index has multiple peaks in the course of each expansion, and the peak

Footnote 5 (continued)
50 percent is the break-even point, this drop in the index implies about a 2 percent decline in manufacturing output. In fact, as the regression estimates predict, output showed no change over this period.

[6] William A. Cox and Theodore S. Torda, "Survey By Purchasing Managers Can Provide Signal On End Of Recession," *Business America*, July 14, 1980, p. 21.

Table 1

"Break-even" Regressions for Manufacturing Growth

Series Explained	Constant	Slope	Predictive Power (R^2) Monthly	Quarterly	Annual	Break-even Point 1959-91	1980-91
Industrial production	−3.62 (11.7)	0.070 (12.9)	.300	.666	.582	51.5	49.3
Payroll employment	−2.44 (17.9)	0.050 (18.4)	.465	.772	.681	49.1	47.4
New orders†	−3.22 (4.5)	0.063 (4.8)	.074	.558	.588	51.4	49.3
Materials inventories	−2.20 (10.2)	0.049 (11.2)	.246	.512	.239	44.7	47.0
Capacity utilization‡	69.24 (89.2)	0.236 (16.9)	.426	.439	.435	—	—
Crude producer prices	−1.89 (3.4)	0.034 (4.1)	.041	.169	.506	56.3	56.1
With the Composite Index							
Industrial production	−3.30 (10.3)	0.067 (11.4)	.250	.556	.669	49.0	46.9
Real GNP	−3.60 (7.4)	0.081 (9.1)	—	.393	.713	44.4	44.5

Notes: Regression coefficients are based on the January 1959–May 1991 sample. Except in the capacity utilization equation, the dependent variable enters as a simple percentage change. Absolute t-values are in parentheses.

†Sample starts in 1967 and the dependent variable is deflated using the implicit deflator for shipments.

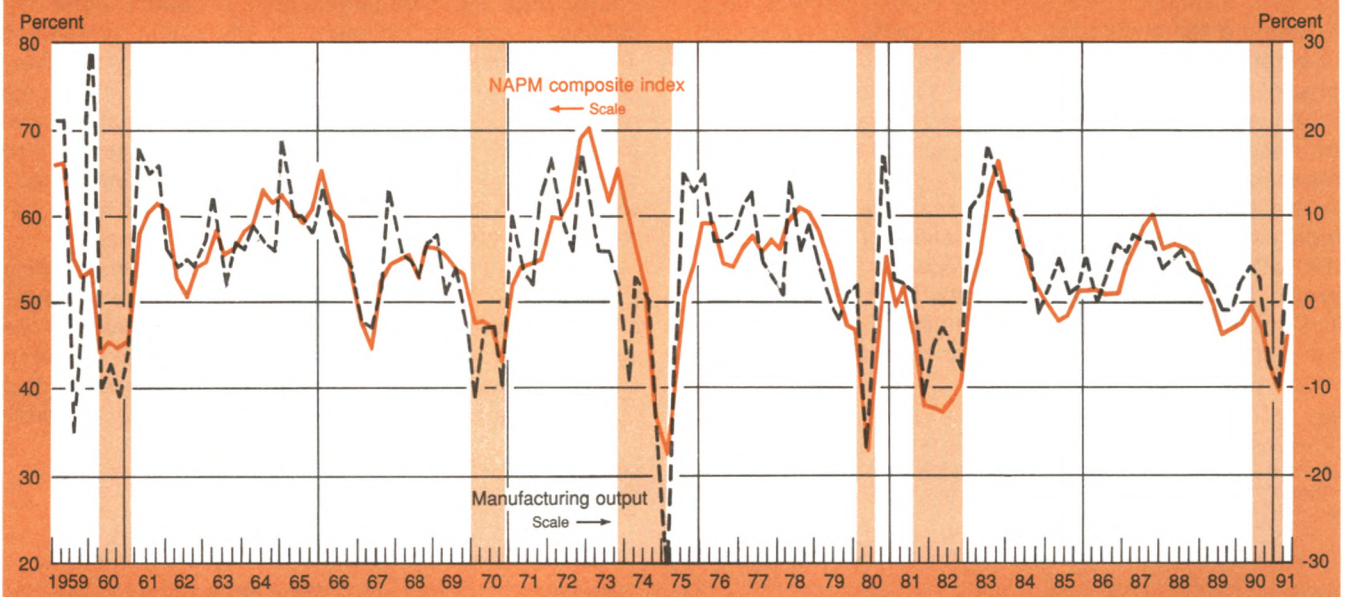
‡The independent variable is vendor deliveries, lagged three months.

that finally "correctly" signals recession can occur anywhere from zero to twenty months before the onset of recessions. The index is just as erratic in predicting cyclical troughs, bottoming out anywhere from zero to twelve months before the economy-wide trough.

If the 50 percent threshold is used rather than the index's peak, equally vexing problems emerge (Table 2). The index usually drops below 50 before cyclical peaks,

Chart 2

The Purchasing Managers' Index and the Growth in Manufacturing Output



Notes: Chart shows quarterly averages of monthly data. Manufacturing output is from the industrial production index. Shaded areas denote recessions.

Table 2

Does the Composite Index Signal Business Cycle Turning Points?

Lead (+) or Lag (−) Time in Months

	NAPM Threshold				NAPM Threshold		
Peak	50.0	49.0	44.5	Through	50.0	49.0	44.5
November 1948	+8	+8	0	October 1949	+1	+1	+2
July 1953	+2	+2	−1	May 1954	0	0	+2
August 1957	+5	+5	−2	April 1958	−2	−2	−1
April 1960	+1	+1	−1	February 1961	−2	−1	−1
December 1969	−1	−1	−9	November 1970	−3	−3	−1
November 1973	−10	−10	−11	March 1975	−5	−5	−3
January 1980	+5	+2	−2	July 1980	−2	−2	−1
July 1981	0	0	−2	November 1982	−3	−3	−2
July 1990	+14	0	−3				
Average	+2.7	+0.8	−3.4	Average	−2.0	−1.9	−0.5
False alarms	4	4	1	False alarms	0	0	1

Notes: In keeping with the leading indicator literature, the composite index is assumed to signal a turning point when it crosses the threshold value for three or more consecutive months. The signal is dated from the first month the threshold is crossed. Signals reversed for at least three months before a cyclical turning point occurs are considered "false alarms."

but the lead time is quite variable. In the 1973-75 recession, the index did not signal recession until almost a year after the onset of the downturn; by contrast, in the most recent recession, the index stumbled along at just below 50 percent for more than a year before the economy turned down. Even worse, it falsely predicted four business cycle peaks, with several signals lasting as much as a year. Its record for cyclical troughs is equally dismal: the index usually surpassed the 50 percent break-even point two to five months after the economy had moved out of recession.

Similar problems arise when threshold values below 50 percent are used. The findings in Table 1 suggest a 49 percent break-even value for industrial production and a 44.5 value for real GNP. Using 49 percent rather than 50 percent as the break-even value has virtually no impact on the timing of the signal. Using 44.5 percent changes the results, but not for the better. At cyclical peaks, the index falls below 44.5 percent after the turning point in all but one downturn, with an average lag of three months. At cyclical troughs, the signal is a little more timely, but again it usually fails to anticipate the recovery. The only advantage of the 44.5 percent threshold is that it produces only two false signals in the postwar period.

Thus the composite index has two problems as a leading indicator. First, because business cycles do not follow a smooth growth pattern, the index often peaks during the initial recovery and then reaches several mini-peaks in the course of an expansion. Second, because growth usually does not flatten out gradually at the peak of the business cycle, the composite index may not dip below 50 percent until after the recession starts. Furthermore, as noted in the appendix, the NAPM data may lag economic activity by about half a month because respondents have incomplete data on the current month when they fill out the survey. Therefore, as a cyclical indicator, the index is better used to *confirm* recent turning points than to anticipate them.

Three horse races

Clearly the composite index and its components have important limitations as stand-alone indicators of the strength of the economy. This section tests how the indexes perform in comparison with alternative forecasting tools. In particular, the analysis explores how the indexes stack up against economic models and consensus forecasts in explaining the growth in nonfarm payroll employment, industrial production, and real GNP. The results confirm that the indexes are poor stand-alone predictors, but they also demonstrate that the indexes provide helpful incremental information to forecasters. In other words, the indexes represent an imperfect but useful addition to our knowledge of current economic conditions.

Forecasting nonfarm employment growth

In the last several years the employment report has become the most important economic indicator for data-watchers.[7] Recognizing this, the purchasing managers' survey committee has pushed up the release date for the index so that it now usually precedes the employment report. Not surprisingly, the composite index and its employment component are viewed as vital information in the payroll employment guessing game.

The explanatory power of the index is tested against two standards. First, the predictions of a simple economic model of employment growth are compared with those of the NAPM data. Second, the performance of the NAPM data is measured against that of the consensus forecast reported by Money Market News Service.

The informal economic model used here is constructed from variables available to forecasters before the purchasing managers' data are released. These include several interest rate spread variables identified in work by Bernanke, Estrella and Hardouvelis, and others as reliable predictors of economic activity.[8] Specifically, the model includes the six-month commercial paper rate, the spread between the commercial paper and Treasury bill rates, the spread between corporate BAA bonds and ten-year Treasuries, and the difference between ten-year and three-month Treasury rates. Also included are several "real" variables watched by payroll forecasters: domestic auto sales, initial claims for unemployment insurance, and the index of leading economic indicators. All told, this ad hoc economic model has eight explanatory variables. The four interest rate variables are entered contemporaneously and with six lags, autos and claims enter currently and with a lag, and both the index of leading indicators and the dependent variable enter with six lags. Adding the NAPM employment index to this model yields a rigorous test of its incremental explanatory power.[9]

Table 3 compares the explanatory power of the economic model, the NAPM employment index, and the full

[7]The markets appear to have a "flavor of the month" approach to economic indicators, with merchandise trade, consumer prices, producer prices, money growth, and the employment report each getting top billing at various times. Overall, however, employment seems to be the most consistent leader.

[8]Ben S. Bernanke, "On the Predictive Power of Interest Rates and Interest Rate Spreads," Federal Reserve Bank of Boston *New England Economic Review*, November-December 1990, pp. 51-68; Arturo Estrella and Gikas A. Hardouvelis, "The Term Structure as a Predictor of Real Economic Activity," Federal Reserve Bank of New York, Research Paper no. 8907, May 1989.

[9]Each model was also tested using the NAPM composite index and using manufacturing employment as the dependent variable, and the results were very similar.

model (that is, the economic model combined with the NAPM index) over two sample periods, one of extended duration (1959-91) and the other limited to recent years (1980-91). For each model the table shows the coefficient on the employment index with its t-statistic and the overall fit of the model as measured by the adjusted R-square. Several results are noteworthy. First, although both the economic model and the purchasing managers' index are highly significant, the economic model explains somewhat more of the variation in employment growth. This finding should not be surprising, however, because the NAPM data measure only growth in the manufacturing sector, while the economic model has a rich array of explanatory variables. Second, and more important, when the NAPM variable is added to the economic model, this variable continues to be highly significant. In fact, the adjusted R-squares suggest that the best model combines the NAPM data and the economic model.[10]

Even stronger support for the NAPM index comes from comparing it with the consensus forecast for payroll employment growth issued by Money Market News Service. This informal survey of data watchers is taken just before the NAPM and employment data are released. The sample is limited to the period since 1985 because of the difficulty in obtaining earlier data. Table 4 shows the results of this comparison. Again, both the NAPM index and the consensus explain a large portion of the variation in employment, but the best results are obtained when the consensus and NAPM are combined in the same equation. This finding suggests that payroll forecasters should modify their forecast in light of the NAPM release. For example, all else equal, a 1 percentage point increase in the NAPM index should induce a 10,000 upward revision in expected payroll employment growth.

So far we have focused on the in-sample fit of the various employment models. The ultimate test of these equations, however, is how they perform out of sample. For each model a series of one-month-ahead forecasts is calculated by using data from 1959 to 1984 and then extending the sample forward one month at a time. Table 5 shows the relative size of the prediction errors for each of the models. As with the in-sample tests, adding either the composite or employment index to the other models reduces the average prediction errors. The best result combines a simple autoregressive

[10] Ideally, it would make sense to modify the estimation in two ways: (1) simplify the model by dropping the less significant lags on each variable and (2) use unrevised data for the independent variables (to duplicate what is available to forecasters). Our purpose here, however, is to stack the odds against the NAPM index as much as possible rather than to devise an optimal model. Furthermore, preliminary tests show that the results are not sensitive to either of these changes.

Table 3
Explaining the Percent Change in Nonfarm Payroll Employment

Model	Sample: 1959-91 NAPM†	$\bar{R}^2$	Sample: 1980-91 NAPM†	$\bar{R}^2$
NAPM index	0.022 (16.9)	.423	0.024 (12.2)	.522
Economic model‡	—	.429	—	.600
Full model§	0.021 (7.3)	.506	0.019 (3.6)	.646

†Values are coefficients on the NAPM index, with absolute t-values in parentheses.

‡Includes the commercial paper rate, three interest rate spread variables, auto sales, initial claims, the index of leading indicators, and lags of the dependent variable.

§Includes both the NAPM employment index and all of the economic variables.

Table 4
Explaining Employment Growth with the NAPM Employment Index and the Consensus Model

Model	Constant	NAPM	Consensus	$\bar{R}^2$
NAPM index	−0.775 (6.0)	0.020 (7.3)	—	.408
Consensus model	−0.035 (1.6)	—	1.173 (10.7)	.600
NAPM combined with consensus model	−0.429 (4.0)	0.009 (3.7)	0.917 (7.5)	.658

Notes: Sample period is January 1985 to May 1991. The dependent variable is the percentage change in total nonfarm employment. Consensus data are converted from change to percentage change. Absolute t-values are in parentheses.

Table 5
Out-of-Sample Prediction Errors for Payroll Employment Growth

Model	Without NAPM	With NAPM: Employment	With NAPM: Composite
NAPM index only	—	.157	.145
Autoregressive model	.143	.141	.136
Economic model	.181	.158	.146

Notes: Table shows the root mean square error for the January 1985–May 1991 period. The "autoregressive model" simply uses six lags on the dependent variable.

model with the composite index.[11]

[11]Note that the economic model alone does the worst job of predicting out of sample for this period. This result is consistent with Bernanke's argument that a structural shift in the relationship between the spread variables and economic activity occurred in the 1980s.

Table 6
Explaining Industrial Production Growth

Model	Constant	NAPM	Consensus	Hours	$\bar{R}^2$
NAPM index	−2.726 (8.4)	0.055 (9.0)	—	—	.371
Consensus model	0.059 (1.3)	—	0.926 (12.8)	—	.544
Hours	0.027 (4.6)	—	—	0.507 (11.4)	.488
Economic model	1.478 (1.6)	—	—	—	.618
NAPM index plus consensus model	−0.401 (0.9)	0.009 (1.1)	0.829 (7.2)	—	.544
NAPM index plus hours	0.178 (4.4)	0.004 (5.1)	—	0.401 (8.8)	.568
NAPM index plus economic model	−0.073 (0.1)	0.030 (2.1)	—	—	.631

Notes: Sample period is January 1980 to May 1991. The dependent variable is the percentage change in the industrial production index. The consensus is from Money Market News Service. Absolute t-values are in parentheses.

Table 7
The NAPM Composite Index and Real GNP Growth

Model	Constant	NAPM	Consensus	$\bar{R}^2$
NAPM index	−14.460 (5.5)	0.326 (6.6)	—	.361
Consensus model	0.546 (0.8)	—	0.832 (4.6)	.217
Economic model	11.346 (4.6)	—	—	.658
NAPM index plus consensus model	−12.000 (4.2)	0.261 (4.5)	0.363 (1.9)	.378
NAPM index plus economic model	0.506 (0.1)	0.152 (1.7)	—	.669

Notes: Sample period is 1970-I to 1989-II. The dependent variable is annualized one-quarter growth in real GNP. Absolute t-values are in parentheses.

Forecasting industrial production and real GNP

The NAPM data are also useful in forecasting industrial production and real GNP. Table 6 compares the explanatory power of four models of industrial production: the NAPM production index, the growth in employee hours, the Money Market consensus forecast, and an economic model using the same variables discussed in the previous section. The t-statistics on the NAPM coefficients suggest that the index adds significantly to the economic model and the simple employee hours model, but that it is not a useful addition to the consensus forecast. This finding should not be a surprise, however, since the NAPM data are available to forecasters before the consensus survey is taken and therefore should already be incorporated into the consensus forecast.

Table 7 shows the results of the final horse race. It compares the power of the composite NAPM index, the economic model, and a consensus forecast to predict growth in real GNP. For the economic model the variables used are the same as those in the employment and industrial production equations, but each variable enters contemporaneously and with two lags. The consensus data, compiled by the American Statistical Association and the National Bureau of Economic Research, are one-quarter-ahead forecasts, taken in the middle of the preceding quarter. Again, the results of the comparison are generally supportive of the NAPM data. The NAPM index predicts real GNP growth better than the consensus forecast, although worse than the economic model. The relatively weak performance of the consensus is easy to explain: the NAPM and economic models use up-to-date information, while the consensus is based only on information available before each quarter. A more important result is that the NAPM index continues to be significant when added to the other models (although it is only marginally significant when combined with the economic model).

Conclusion

Despite its growing popularity, the NAPM index has undergone very little critical scrutiny. Our results suggest that the index is a flawed but still useful indicator. It is a poor leading indicator and, on its own, can be a misleading measure of short-run movements in the economy. In combination with other data, however, it is very helpful in predicting contemporaneous manufacturing activity. In sum, the index deserves at least part of its reputation as a key economic indicator.

Appendix: The Design of the NAPM Data Set

With one notable exception, the NAPM data have received high praise in the literature.[†] Hoagland and Taylor, for example, argue that the survey data "are available sooner, are more reliable, and are much more cost effective than government information."[‡] Klein and Moore cite the early release of the data as an important advantage; they recommend that the inventory index be substituted for the official inventory data to improve the timing of the index of leading indicators.[§]

Despite this strong support, the NAPM data need improvement in at least three important areas.

Sampling bias

Unlike the surveys underlying official statistics, the purchasing managers' survey does not use a scientific sample. The NAPM data are drawn from hand-picked members of larger, older firms rather than from a probability sample. No attempt is made to account for industry growth through the increase in the number of firms. Furthermore, newer, fast-growing firms are added to the sample only after they have become established in the business, while declining firms remain in the sample until they go out of business. In official statistics, both of these downward biases are eliminated through adjustments and rebenchmarking.

The sampling design has additional problems. The sample is small, comprising less than 1 percent of the association's membership. Because of nonresponses and the entry and exit of members, firms answering the survey questionnaire can vary from sample to sample. No attempt is made to correct for this variation by linking companies that respond in both the current and previous month—a procedure followed in the official statistics. Finally, the data are never revised, implying that errors are never corrected and late responses are never incorporated into the data.

These sampling problems may explain the apparent downward bias in the indexes. Theoretically, when aggregate activity is unchanged, the indexes should read about 50 percent, with equal numbers of firms reporting higher and lower activity. In fact, as Table 1 in the text shows, the break-even values tend to be well below 50 percent. The results for the inventory index are particularly troubling. Not only is the break-even point well below 50 percent, but the index also *averages* only 47.8 percent over the entire postwar period. This finding implies that the level of inventories held by manufacturing firms has had a downward trend. Government statistics, on the other hand, show inflation-adjusted materials and supplies for manufacturers roughly doubling over this period.[‖]

Backward-looking data

An important attribute of the NAPM data—its timeliness—is also one of its biggest shortcomings. Since the results are released just after the end of each month, the questionnaire must be answered in the middle of the month. As a result, when respondents try to compare the "current" month with the "previous" month, they may in fact be comparing their impression of the last few weeks (including part of the previous month) with their recollections of the weeks before that interval. As the table below shows, the timing of the responses means that in some cases the NAPM data are more closely correlated with lagged activity than with current activity.

Subjective responses

Survey respondents may not accurately assess whether conditions are "better" or "worse." Their answers may reflect what should be or what is projected rather than what is. The low average reading for inventories, for

[†]The exception is Feliks Tamm, "An Agenda for Inventories Input to the Leading Composite Index," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators*, (Cambridge: Cambridge University Press, 1991), pp. 429-60. Tamm points out a variety of flaws in the NAPM inventory data. Some of his concerns are discussed here.

[‡]John H. Hoagland and Barbara E. Taylor, "Purchasing Business Surveys: Uses and Improvements," *Freedom of Choice: Presentations from the 72nd Annual International Purchasing Conference* (Oradell, N.J.: National Association of Purchasing Management, 1987), p. 1.

[§]Philip A. Klein and Geoffrey H. Moore, "N.A.P.M. Business Survey Data: Their Value as Leading Indicators," *Journal of Purchasing and Materials Management*, Winter 1988, pp. 32-40.

[‖]Comparing the NAPM indexes for employment and output with the official diffusion indexes for manufacturing employment and industrial production confirms this bias. Regression estimates for the 1980-91 period show that the break-even values for both official diffusion series are closer to 50 percent.

Correlation of NAPM Indexes and Manufacturing Data

Official Series	Lead	Contemporary	Lag
Industrial production	.410	.547	.614*
Employment	.569	.682	.720*
New orders	.211	.272	.426*
Materials inventories	.481	.496*	.476

Notes: The sample period is January 1959 to May 1991. The asterisk indicates peak correlation.

Appendix: The Design of the NAPM Data Set *(continued)*

example, may reflect the constant concern about excessive stocks rather than actual inventory management. Wishful responses are particularly likely since the sample is taken before the full results for the month are known, and many of the questions refer to areas of the firm not under the direct purview of the purchasing manager.

The response that economic activity is "the same" is equally problematic. Over time an average of more than half the responses is "the same." For example, from January 1990 to June 1991 the percentage of "same" responses was: new orders (46.5), production (53.9), inventories (53.1), vendor deliveries (82.3), and employment (64.4). Such stability at the firm level seems quite unlikely in an unstable period for the economy as a whole. Apparently, "the same" is a catch-all assessment meaning "don't know" and "no response" as well as "no change."

Improving the data

In a real sense the NAPM data set is an uncut gem. By using modern sampling and statistical techniques, the association could greatly improve the accuracy of the data. A probability sample should replace the hand-picked sample; respondents should be linked from one sample to the next; efforts should be made to reduce the number of "same" responses and to ensure that responses reflect actual activity; and respondents should be encouraged to report on the current month's activity only. In addition, correctly accounting for inventories and adjusting for lags and leads in the components would improve the composite index.[††] Of course, the NAPM data neither could nor should mimic the official statistics: this would require delays in its release and would put an impossible burden on the respondents. The purchasing managers' association has made some efforts to refine the data. Nevertheless, with the index increasingly in the spotlight, further modernization is warranted.

[††]The inventory index should enter the composite index as a first difference rather than a level since it measures a stock rather than a flow. In a forthcoming paper, Mark Flaherty and the author present an alternative composite index that has an improved track record in predicting industrial production, real GNP, and the index of coincident indicators.