A policymakers' guide to economic forecasts

Steven Strongin and Paula S. Binkley

Economic forecasts have always had an uneasy relationship with monetary policy. Policy clearly must focus on the future; however, just as clearly, policymakers need a firmer foundation for current actions than projections about tomorrow's economic performance. The reluctance of policymakers to rely on economic forecasts is understandable. There is a large body of research evaluating the accuracy of economic forecasts and demonstrating that, while forecasts do contain useful information, they can also be seriously misleading.

To offset these well-known limitations of economic forecasts, policymakers have used historically robust long-run relationships to supplement economic forecasts. The targeting of the money supply from 1979 to 1982 is a good example of a case where the long-run relationships were dominant over short-run forecasts in the formation of policy.

In the 1980s we have seen many of these long-run relationships falter, for a variety of reasons, including financial deregulation and industrial restructuring. Many economic relationships of great historical persistence have simply vanished. For example, velocity, the key summary measure of the usefulness of monetary targeting, has been extremely unstable during this decade. As a result of the failures of these long-run guides to policy, forecasts have become increasingly important in the process of policy formation.

If policy is going to be forced to rely more heavily on economic forecasts, it is useful not only to analyze the general accuracy of those forecasts, as previous studies have done, but also to assess the accuracy of revisions to those forecasts. The making of policy is a continuous process. New information arrives, that information is incorporated into the policymaker's understanding of the economy, and then policy is revised in accordance with that revised understanding.

Unfortunately, economic information is not always accurate. Often new information that seems very important turns out to be either a short-run anomaly or just plain wrong, later revised to show an entirely different course of events. As a result, forecasts are sometimes revised in the wrong direction and actually become less accurate as new information is applied.

The frequency with which new information is misleading has strong implications for the day-to-day management of policy. If incoming information is highly reliable then policy should respond quickly and forthrightly to that information. On the other hand, if the reliability of incoming information is in serious question then the policymaker should wait for confirming evidence. In general, the lower the reliability of the new information, the more evidence the policymaker should require before changing policy.

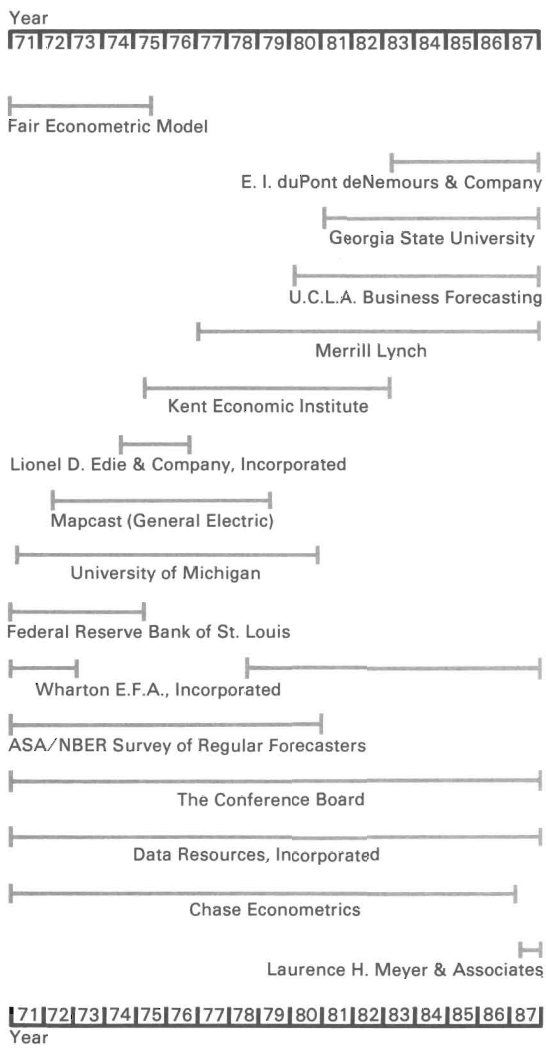
This study analyzes the accuracy of forecast revisions in order to determine the amount of evidence that should, in general, be required to significantly improve the policymaker's information about the current economic outlook. The results of the analysis are dramatic. Forecast revisions, *during the year being forecast,* are as likely to worsen as to improve the accuracy of the forecast with as much as six months' additional economic information. These results make a strong case for policymakers or any other users of economic forecasts to be extremely cautious when responding to current economic statistics. We found that forecasts were often revised significantly in the *wrong* direction in response to supposedly clear economic signals. Early forecasts were often more accurate than their mid-term forecast revisions. In such an environment, it is important for policymakers to wait for full verification of economic trends before acting. Early signals simply are not reliable.

Methodology

This study uses a straightforward approach to analyzing revisions to economic forecasts. The Conference Board publishes, in its monthly *Statistical Bulletin*, a set of current, publicly available, economic forecasts. These

Steven Strongin is a senior economist and assistant vice president, and Paula S. Binkley, an associate economist at the Federal Reserve Bank of Chicago.

Figure 1
Forecasters and periods of publication



forecasts were used to construct a small database of revisions to macroeconomic forecasts, specifically forecasts of real GNP growth and inflation. Figure 1 shows the forecasters whose forecasts were published and over what time periods those forecasts were available. There was a wide range of participants from academics, such as UCLA Business Forecasting, to commercial forecasting firms, such as DRI and Chase Econometrics.

While some other broader surveys, such as the Blue Chip, could have been used, the Conference Board survey offered a number of advantages. All of the forecasts it contains are normally publicly available and the methods of the participants well known. The relatively limited number of participants allowed us to examine the numbers closely to make sure that no peculiarity of some individual forecast or small group was generating the results. Further, the sample is monthly back to 1971 which produced a good balance between frequency of forecasts and years of availability.

The forecasts for any given month were averaged to create a single forecast. The averaging was used to represent the overall evolution of a consensus view of the economy. From a purely statistical viewpoint, other methods may be preferable, but a simple average probably captures the essence of how forecasts enter the policy process. The study does not attempt to model or examine whether or not forecasters are using information correctly; it restricts its attention to the overall flow of information which the policymaker receives. Specifics of the forecasting process are ignored in favor of analyzing the final output of the forecasting process.

It might be argued that it would have been more informative to examine the actual data streams of incoming information. This approach suffers from two significant flaws. The first is that different information sets are given different weights at different times. Economic series are often contaminated by widely known events. For instance, capacity utilization numbers are often discounted because of ongoing structural adjustments in the economy. Financial flow variables are often discounted around April 15 because of the well-known but nevertheless random effects of tax day. The use of forecasts, rather than actual data streams, abstracts from the difficulties of having to adjust for known, but statistically unstable, flaws in the data.

The second problem in examining the specific data flows is also the more fundamental. From a policy perspective, the question is not why forecast revisions have a given behavior pattern, it is what information can be derived from those revisions. Whether the poor performance of forecast revisions that we will examine later in this paper is due to the quality of information, the way that information is used, or some other peculiarity of the process, is secondary to an understanding of exactly how bad that performance is. As we attempt

to address the problems that the current study raises it will be necessary to examine the why's, but that is beyond the scope of this paper.

Revisions were studied over the year being forecast. For instance, when we discuss the forecasts for 1987, the January forecast of 1987 would be the forecast of growth from fourth quarter of 1986 through the fourth quarter of 1987 made in January of 1987. And similarly, the December forecast would be the forecast of growth from the fourth quarter of 1986 to the fourth quarter of 1987 made in December of 1987. In a very real sense, the December forecast would be a forecast of the previous year. Thus, as the year progresses the information being built up is not only about trends, but about actual economic performance of the period being forecast. In such a case, it might seem reasonable to assume that a rather steady and rapid improvement in forecasts would be evident. As we shall see, this is not the case.

Forecasts of real GNP growth

The time paths of each year's forecast errors for real GNP growth and inflation are shown in the graphs in Figure 2. The graphs show the error in each forecast as measured against the first Bureau of Economic Analysis benchmark estimates of actual economic performance. As can be seen, there is a tremendous diversity of experience. Only 1982 exhibits what might be naively thought of as a typical pattern of learning—a year either stronger or weaker than expected, with forecasters steadily updating their forecasts in the appropriate direction. More typical of actual experience, though clearly the word typical overstates the case, is 1987, where the initial forecasts were fairly accurate. As the first few revisions were made the forecast drifted off, and then about mid-year the forecast began to steadily improve.

It is evident from an examination of the patterns of revisions in the graphs contained in Figure 2 that it is very common for forecasts to drift away from, rather than toward, truth. There is a very good reason for this. As the forecasts originally had all previous information built into them, revisions to forecasts are totally at the mercy of the very latest economic statistics and thus likely to reflect all of the limitations and false signals contained in those statistics. So the fact that forecasts often track off-course does not come as any particular surprise. It just demonstrates what policymakers and forecasters have known for a long time—that great caution should be exercised when forecasting economic trends from one month of data.

The key question is how much data is necessary to give the policymaker some confidence that the latest numbers are actual evidence of a trend rather than statistical noise. The magnitude of the difficulties in interpreting incoming data comes out quite strikingly in Table 1. For the sake of simplicity the number of forecasts examined is reduced to January and the end of each of the four quarters of the year. Table 1 compares the accuracy of forecasts at the end of each quarter to the accuracy of the forecasts made in January. This is useful because in January no actual data for the year being forecast is available, but as time goes on every quarter's forecast contains an additional quarter's worth of economic information about the year being forecast.

In only 7 of the 17 years covered by the database was there any improvement in the forecast of real GNP by the end of the first quarter and in only 8 of the 17 years was there any improvement by the end of the second quarter. This is less than an even money result, indicating that through the first half of the year current economic information is at least as likely to mislead the economic forecaster or policymaker as it is to help. By the end of the third quarter, in 11 of the 17 years the forecast has moved in the right direction. This is at least indicative of some improvement; though on purely statistical grounds, it is still insufficient to flatly reject the hypothesis that incoming information is not providing any useful input to the forecaster.[1] By the end of the fourth quarter, forecast accuracy has improved substantially.

Table 1
Real GNP forecast improvement record

Periods compared	Number of years in which improvements occurred
Q1 to January	7 out of 17
Q2 to January	8 out of 17
Q3 to January	11 out of 17
Q4 to January	15 out of 17

Figure 2
Forecast error time paths

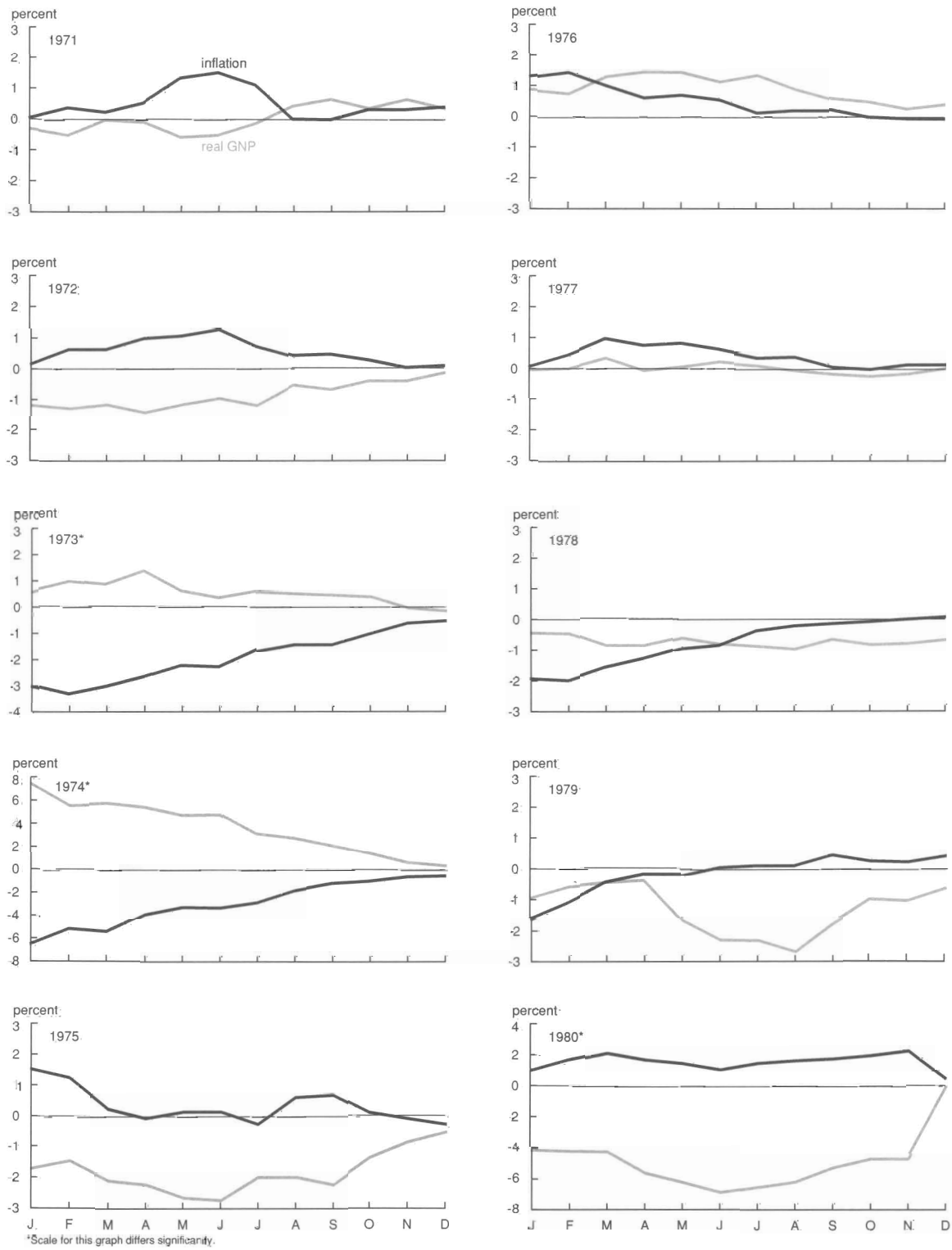
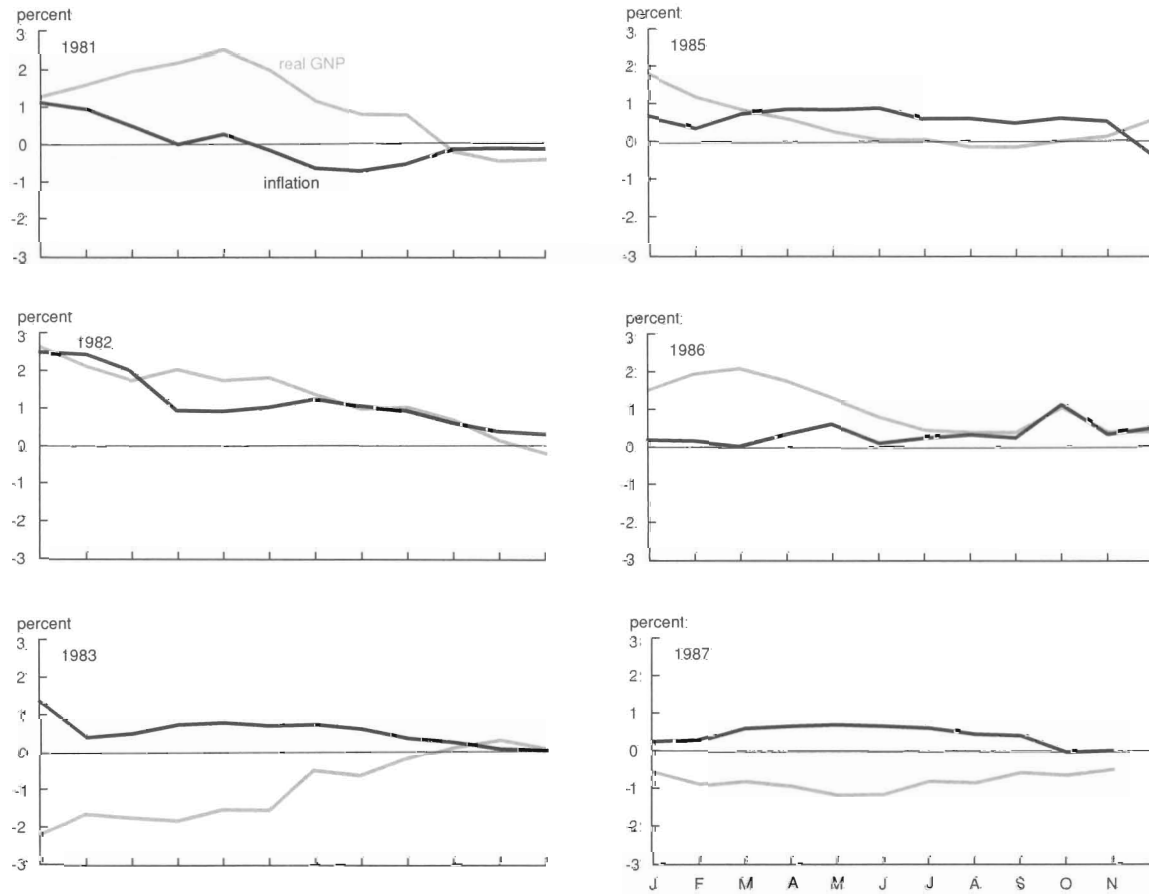


Figure 2
Forecast error time paths (cont'd)



One immediate question that the very poor improvement record of real GNP forecasts raises is whether this a peculiarity of some particular set of years. Table 2 shows the percentages of improvement in forecasts for the first and second half of the time period investigated. While the 1971-1979 period is worse than the 1980-1987 period, neither period shows substantially different qualitative results, especially considering the relatively small span of years covered by the sub-samples. In both sub-periods, it is still not possible to conclude with any confidence that there is a substantial improvement in forecast accuracy until the end of the third quarter.

Another way of looking at the data is in terms of mean absolute error in the forecasts at different points of the year. Put more sim-

Table 2.
Real GNP forecast improvement record*

Periods compared	1971-1986	1971-1979	1980-1987
Q1 to January	41%	33%	50%
Q2 to January	47%	33%	63%
Q3 to January	65%	44%	88%
Q4 to January	88%	76%	100%

*Percent of forecast revisions which improve forecast

Table 3
Magnitude of difference between forecasted and actual real GNP growth rates

Forecast as of	Average absolute error
January	1.70
End of Q1	1.59
End of Q2	1.67
End of Q3	1.05
End of Q4	.35

ply, what is the average size of forecast error across the year being forecast? This is subject to some objections because, unlike statistics based on the percentage of years in which forecasts improve, statistics about size of error can easily be dominated by a few especially bad years. Nevertheless, Table 3 shows the mean absolute error of forecasts at the same point of time during the years investigated in previous tables. Here the picture looks somewhat better, with some moderate improvement over the year, though even here the forecasts actually get worse from the end of the first quarter to the end of the second quarter.

However, the graphs in Figure 2 indicate that much of the observed improvement in forecast accuracy could be due to rather massive improvements in forecast accuracy in 1974 and 1980 when the January forecasts were very wrong (so much so, in fact, that those years required special scaling in Figure 2). Table 4 shows the results of Table 3 when 1974 and then 1974 and 1980 are removed from the data sample. With the exclusion of these years no

Table 4
Magnitude of difference between forecasted and actual real GNP growth rates

Forecast as of	Average absolute error
(Excluding 1974)	
January	1.34
End of Q1	1.34
End of Q2	1.48
End of Q3	.99
End of Q4	.35
(Excluding 1974 and 1980)	
January	1.15
End of Q1	1.14
End of Q2	1.12
End of Q3	.70
End of Q4	.37

improvement in forecast accuracy is noticeable until the end of the third quarter. So, at least for real GNP forecasts, the data would indicate that six to nine months of data would normally be required to provide a good case for revising policy based on economic forecasts. Even a very sympathetic reading of the data could not push that number much lower than six months. This is not to imply that large-scale events should not cause policymakers to act, but it does suggest that policy actions should be based on a fundamental understanding of the events themselves and not merely derived from incoming, and often preliminary, economic statistics or economic forecasts.

One important thing to remember in this regard is that an event that might cause a recession in the absence of a policy change may not do so if policy does, in fact, react. In such a case economic forecasts might go wrong not because they fail to understand the economic events, but because they underestimate the policymaker's ability or willingness to respond to those events. Thus, the statistics presented above cannot be used to argue against policy action, but only provide a cautionary note about the amount and type of information required to justify policy action.

Forecasts of inflation

Here, the news is a little better. Forecasts of inflation show a marked tendency to improve with additional information. As Table 5 shows, at the end of the first quarter there is an improvement in 9 of 17 years, and by the end of the second quarter there is improvement in 11 of 17 quarters, indicating that within 6 months there is some reason to think that inflation forecasts have improved. Table 6 shows the time sub-sample breakdown of this result. The sub-sample results appear to indicate that much of the observed improvement in the accuracy of inflation forecasts was due to the

Table 5
Inflation forecast improvement record

Periods compared	Number of years in which improvements occurred
Q1 to January	9 out of 17
Q2 to January	11 out of 17
Q3 to January	12 out of 17
Q4 to January	14 out of 17

Table 6
Inflation forecast improvement record*

Periods compared	1971-1986	1971-1979	1980-1987
Q1 to January	53%	56%	50%
Q2 to January	65%	67%	63%
Q3 to January	71%	78%	63%
Q4 to January	82%	78%	88%

*Percent of forecast revisions which improve forecast

1971-1979 period, with the more recent period again showing no real evidence of improvement until the end of the third quarter. However, an examination of the graphs in Figure 2 shows that many of the mistaken revisions to inflation forecasts in the 1980-1987 time period occurred in years in which the inflation forecast started out very accurate and varied little throughout the year, though veering slightly off. In fact, careful analysis of the graphs of the 1980s indicates that on the whole inflation forecasts do improve throughout the year.

In support of this visual evidence, Table 7 shows the average absolute error of inflation forecasts throughout the year and Table 8 shows the time sub-sample results. These numbers do indeed show a consistent improvement in the overall accuracy of inflation forecasts throughout the year. It should also be noted, in comparing Tables 7 and 8 to Tables 4 and 5 for real GNP growth, that inflation forecasts are on the whole more accurate and improve faster and with more consistency than the real GNP growth forecasts.

Table 7
Magnitude of difference between forecasted and actual inflation growth rates

Forecast as of	Average absolute error
January	1.45
End of Q1	1.27
End of Q2	.93
End of Q3	.56
End of Q4	.26

A digression

A side note is in order at this point. The view of forecasting and policy presented here is not universal. The criticism has often been made of both policymakers and forecasters that

Table 8
Comparisons of inflation forecasts over different periods

Period	Average absolute error				
	January	End of Q1	End of Q2	End of Q3	End of Q4
1971-1987	1.45	1.27	.93	.56	.26
1971-1979	1.80	1.51	1.18	.52	.28
1980-1987	1.05	1.00	.64	.60	.25

they are not sufficiently sensitive to turning points in the economy and that as a result the economy has, all too often, been allowed to slip into recession. In this view, it is all right if we occasionally respond to weakness in the economy that is not there, so long as we respond to weakness when it is there.

The problem with this view is that it substantially underestimates the practical limitations of economic statistics implied by the previous analysis. While the specifics of forecasting turning points were not studied, it seems unlikely that economic data that cannot reliably be used to revise a forecast would nevertheless be sufficient to sniff out a sudden unexpected recession. Rather, it seems likely that if policy strongly reacted to downward revisions in the outlook it would be in a constant fight against imaginary recessions, inevitably leading to major inflationary excesses.

The economy generates too many false signals to allow reliable forecasting of events that evolve quickly, such as economic turning points. Thus, at least within the context of normal forecasting methods and our current ability to interpret economic information, turning points are unlikely to provide a reasonable focus for policy.

Conclusions

From the point of view of the day-to-day management of policy, it is clear that changes in the economic outlook provide only minimal guidance to policymakers. Further, even over the span of six to nine months more attention should be paid to revisions in inflation forecasts than to revisions in real GNP growth forecasts. This is not to say that economic analysis provides little benefit to policymakers. It simply points out the fact that it is the analysis of real events and fundamental factors in the economy that must guide policy and not the last round of forecasts based on the very latest economic statistics.

These results might be considered a guide to what we can reasonably expect economic analysis to provide to the policy process. It can provide a baseline understanding of how the economy is operating and how the overall thrust of policy fits into that operation. What we cannot expect from economic analysis and should not expect from policy is real time management of the economy. The data and our ways of filtering that data are simply not up to the task.

[1] Technically, if the direction of forecasts to or away from truth is modeled as a binomial process in 17 trials, 12 correct revisions would be required to reject the null hypothesis of incoming economic information containing no useful information at the 95% confidence level.

Appendix

Difference between forecasted and actual fourth quarter to fourth quarter growth rates

Forecast year	Difference (percentage points) January	End of Q1	End of Q2	End of Q3	End of Q4
	Real GNP				
1971	−0.30	−0.01	−0.50	0.64	0.36
1972	−1.18	−1.18	−0.98	−0.68	−0.15
*1973	0.58	0.83	0.33	0.46	−0.16
1974	7.42	5.64	4.75	2.03	0.32
*1975	−1.73	−2.10	−2.73	−2.23	−0.55
1976	0.90	1.34	1.17	0.59	0.40
1977	−0.03	0.33	0.22	0.16	0.02
1978	−0.43	−0.87	−0.80	−0.66	−0.67
1979	−0.96	−0.46	−2.30	−1.75	−0.64
*1980	−4.14	−4.26	−6.89	−5.28	−0.09
*1981	1.30	1.97	1.98	0.73	−0.42
*1982	2.64	1.73	1.80	0.99	−0.21
1983	−2.18	−1.75	−1.58	−0.19	0.08
1984	−1.23	−0.86	−0.41	0.35	−0.39
1985	1.79	0.89	0.07	−0.13	0.57
1986	1.51	2.07	0.82	0.41	0.41
1987	−0.56	−0.78	−1.13	−0.55	−0.47**

Forecast year	Difference (percentage points) January	End of Q1	End of Q2	End of Q3	End of Q4
	Inflation				
1971	0.06	0.23	1.50	0.00	0.41
1972	0.17	0.64	1.28	0.42	0.08
1973	−3.01	−3.05	−2.31	−1.43	−0.51
1974	−6.49	−5.41	−3.28	−1.18	−0.51
1975	1.53	0.25	0.16	0.69	−0.29
1976	1.32	1.04	0.58	0.23	−0.04
1977	0.07	0.96	0.63	0.07	0.15
1978	−1.92	−1.57	−0.84	−0.14	0.07
1979	−1.62	−0.44	0.07	0.48	0.45
*1980	0.99	2.20	1.11	1.82	0.47
*1981	1.13	0.49	−0.16	−0.54	−0.16
*1982	2.50	1.98	1.03	0.92	0.30
1983	1.36	0.52	0.71	0.36	0.02
1984	1.32	1.32	0.36	0	0.09
1985	0.68	0.77	0.92	0.48	−0.37
1986	0.20	0.04	0.13	0.27	0.53
1987	0.25	0.64	0.69	0.44	0.04**

*Years in which NBER turning points occurred.
**Figure for November 1987, last available forecast.