When can we forecast inflation?

Jonas D. M. Fisher, Chin Te Liu, and Ruilin Zhou

Introduction and summary

The practice of forecasting inflation has generally been considered an important input in monetary policymaking. Recently, this view has come under attack. In an article that appeared in the Federal Reserve Bank of Minneapolis's *Quarterly Review*, Atkeson and Ohanian (2001, hereafter A&O) argue that the likelihood of accurately predicting a change in inflation using modern inflation forecasting models is no better than a coin flip. They conclude that these forecasting models cannot be considered a useful guide for monetary policy. In this article, we reexamine the findings that underlie this conclusion. We show that it may be possible to forecast inflation over some horizons and in some periods.

A&O study the properties of standard *Phillips-curve-based* inflation forecasting models. These models relate changes in inflation to past values of the unemployment gap (the difference between unemployment and a measure of unemployment believed to be associated with non-accelerating inflation, the so-called NAIRU [non-accelerating inflation rate of unemployment]), past changes in inflation, and perhaps other variables believed to be useful indicators of inflation.[1] Recently, Stock and Watson (1999, hereafter S&W) proposed a generalized version of the Phillips curve and argued that their generalization is superior to these standard models as a forecasting tool. Focusing on the one-year-ahead forecast horizon, A&O argue that unemployment-based Phillips curve models and S&W generalized Phillips curve models can do no better than a "naive model," which says that inflation over the coming year is expected to be the same as inflation over the past year. This analysis focuses on the ability to forecast the magnitude of inflation in the Consumer Price Index (total CPI), the CPI less food and energy components (core CPI), and the personal consumption expenditures deflator (total PCE) over the sample period 1985 to 2000.

To gain some insight into these findings, figure 1, panel A displays 12-month changes in 12-month core CPI from 1967 to 2000. The vertical lines in this figure (in 1977, 1985, and 1993) divide the sample period into four periods. It is immediately clear that in the two later periods, that is, the sample period considered by A&O, the volatility of changes in inflation was much lower than in the two earlier periods. This change in the behavior of inflation seems to be coincident with the change in monetary policy regime that is generally thought to have taken effect in the mid-1980s.[2] The lower volatility and the possibility of a changed monetary policy regime in the later two sample periods may favor the naive model studied by A&O. Figure 1, panel B shows that PCE less food and energy components (core PCE) behaves in a similar fashion.

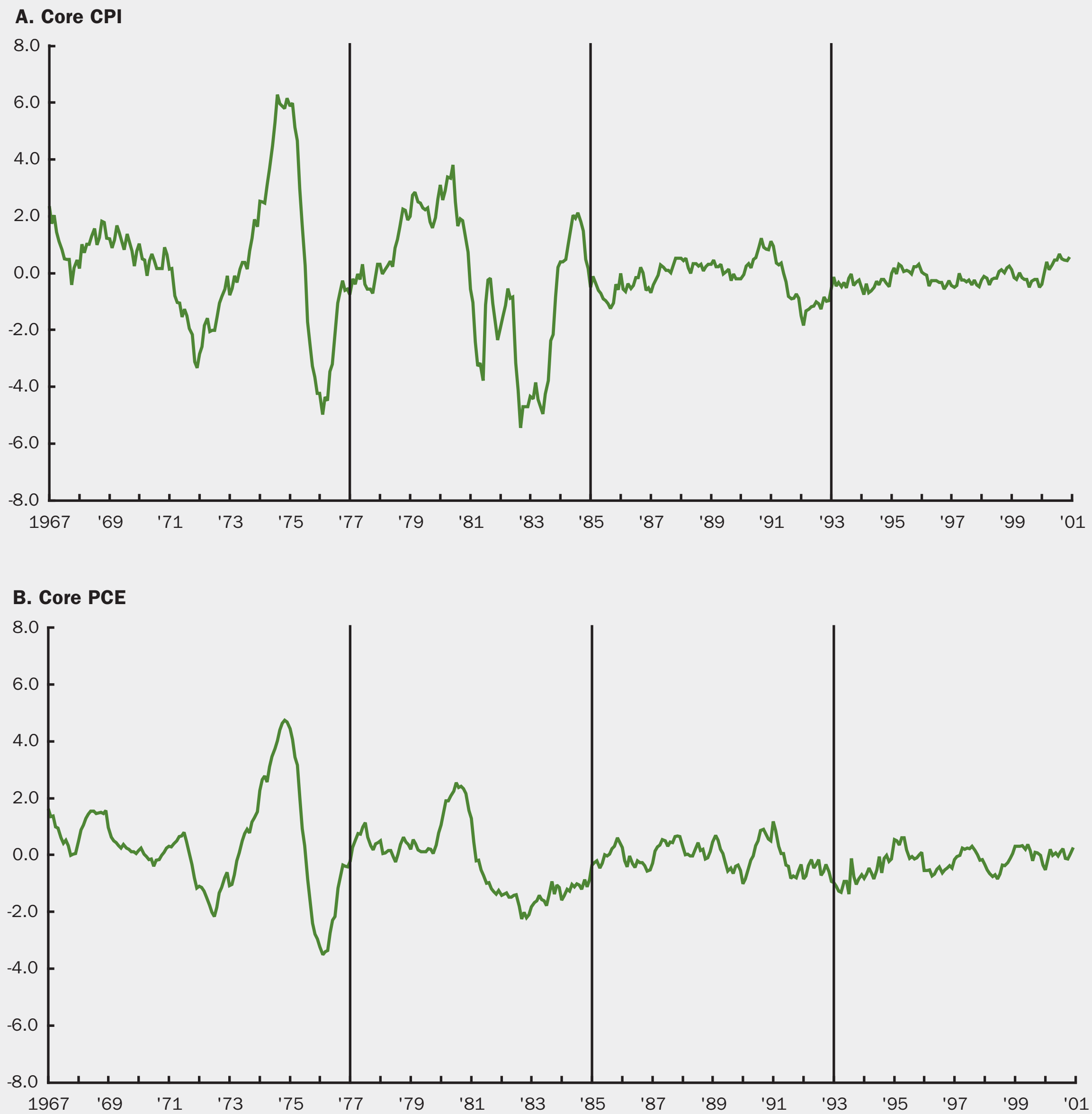
These changes in the behavior of inflation raise the question of whether A&O's findings are due to special features of the data in the sample period they chose to focus on. To address this possibility, we extend the A&O analysis by studying three distinct sample periods, 1977–84, 1985–92, and 1993–2000. In addition, we add core PCE inflation to the list of inflation measures and we consider a broader class of Stock–Watson type models. A&O focus on the one-year forecast horizon. Given the lags inherent in the effects of monetary policy actions, it is reasonable to consider whether their results extend to longer horizons. Consequently, we analyze both the one-year and two-year forecast horizons.

Our findings confirm the A&O results for the 1985–2000 period, but not for 1977–84. The Phillips curve models perform poorly in both the 1985–92 and 1993–2000 periods when forecasting core CPI.

Jonas D. M. Fisher is a senior economist, Chin Te Liu is an associate economist, and Ruilin Zhou is a senior economist at the Federal Reserve Bank of Chicago.

FIGURE 1

12-month changes in 12-month core inflation



Source: Haver Analytics, Inc., 2001, U.S. economic statistics database, July.

However, when forecasting core PCE, these models improve significantly relative to the naive model in the 1993–2000 period. While the Phillips curve models do poorly for the one-year-ahead forecast horizon, we do find evidence in favor of the Phillips curve models for the two-year-ahead forecast horizon, at least with respect to core inflation. Taken together, these findings are consistent with our suspicion that periods of low inflation volatility and periods after regime shifts favor the naive model.

The relatively poor performance of the Phillips curve models reflects their inability to forecast the magnitude of inflation accurately. Ultimately, the way we assess our forecasting models should reflect the usefulness of the forecasts in policymaking. In our view, policymakers understand that precise forecasts of inflation are fraught with error. As a result, they pay considerable attention to the direction of change of future inflation. For this reason, we do not view measures of forecast performance used by A&O and many others that emphasize magnitude as the only criteria for evaluating forecasting models.

Consequently, we consider a complementary approach to evaluating forecasting models that emphasizes the forecasted direction of change of future inflation. Under the assumption that forecast errors are symmetrically distributed about the forecast, the naive model provides no information about future inflation; it is no better than a coin flip at predicting the future direction of inflation. Under the same symmetry assumption, the Phillips curve models predict that inflation will change in the direction indicated by comparing the point forecast with the current level of inflation. We analyze the ability of our Phillips curve models to forecast the direction of inflation and find that they do quite well. Over the entire 1977–2000 period, the Phillips curve models are able to forecast the correct direction of inflation one year ahead between 60 percent and 70 percent of the time. For the same period, the models forecast the correct direction two years ahead more than 70 percent of the time.

These results suggest that the Phillips curve models forecast the direction of inflation changes relatively well across measures of inflation and across time. But when it comes to forecasting the magnitude of inflation changes, there may be times, such as after a change in monetary policy regime, when the naive model may do better than the Phillips curve models. The last question we address is whether it is possible to improve on the forecasts of the naive model in difficult times by using the directional information contained in the Phillips curve models. We show that it is possible to improve on the naive model, although the improvement is modest.

One interpretation of our findings is that it is possible to forecast inflation accurately during some periods, but not others. We argue that the periods in which it is difficult to forecast inflation are associated with changes in monetary policy regime, broadly interpreted. This implies that if we are in a stable monetary regime and expect the regime to persist, then it may make sense for policymakers to pay attention to inflation forecasts.

In the next section, we outline the different forecasting models that we consider in our analysis. Next, we discuss the standard methodology we implement to evaluate the ability of these models to forecast the magnitude of future inflation. We then discuss our results for forecasting magnitude and present our analysis of forecasting directional changes in inflation. We describe our procedure for combining the naive model with our directional forecasts and how well this procedure performs over our sample period. Finally, we discuss some possible policy implications of our findings.

Statistical models of inflation

The standard approach to forecasting inflation is rooted in ideas associated with the Phillips curve, the statistical relationship between changes in inflation and measures of overall economic activity. The generalized version of the Phillips curve proposed by S&W involves variables that summarize the information in a large number of inflation indicators. S&W argue that their generalization is superior to conventional Phillips curves as a forecasting tool. A&O argue that neither the conventional nor the generalized Phillips curve framework can do better than a simple forecast (their naive model) that says inflation over the coming year is expected to be the same as inflation over the past year. We reexamine this claim using a broader class of S&W-type models than considered by A&O. Now we describe in detail the models we study.

The naive model

The benchmark for evaluating our models is the naive model described by A&O. The starting point for the naive model is the *martingale hypothesis*, which states that the expected value of inflation over the next 12 months is equal to inflation over the previous 12 months. Specifically,

$$1)\quad E_t\pi^{12}_{t+12} = \pi^{12}_t,$$

where the 12-month inflation rate, π^{12}_t is defined as the 12-month change in the natural logarithm of the price indexes p_t,

$$\pi^{12}_t = \ln p_t - \ln p_{t-12},$$

and E_t denotes the expectation conditional on date t information. The naive model equates the forecast of inflation over the next 12 months, $\hat{\pi}^{12}_{t+12}$, with its conditional expectation. That is,

$$2)\quad \hat{\pi}^{12}_{t+12} = \pi^{12}_t.$$

Notice that if the martingale hypothesis holds, then the expected value of 12-month inflation in the second year following date t must also equal inflation over the 12 months prior to date t, that is

$$E_t\pi^{12}_{t+24} = \pi^{12}_t.$$

Similar to the 12-month forecast, the naive model equates the forecast of inflation over the next 24 months, $\hat{\pi}^{12}_{t+24}$, with its conditional expectation:

3) $\hat{\pi}^{12}_{t+24} = \pi^{12}_t$.

Generalized Phillips curve models

The simplest alternative to the naive model postulates that changes in 12-month inflation only depend on recent changes in one-month inflation. That is, for $J = 12, 24$,

$$4) \quad \pi^{12}_{t+J} - \pi^{12}_t = \alpha + \beta(L)\,(\pi_t - \pi_{t-1}) + \varepsilon_{t+J}\,,$$

where the one-month inflation rate, π_t, is defined by

$$\pi_t = \ln p_t - \ln p_{t-1}.$$

In addition, ε_t is an error term, and $\beta(L)$ specifies the number of lags in the equation.[3] Below, we refer to this as model 1.

The next model we consider is based on the Chicago Fed National Activity Index (CFNAI). This index is a weighted average of 85 monthly indicators of real economic activity. The CFNAI provides a single, summary measure of a common factor in these national economic data. As such, historical movements in the CFNAI closely track periods of economic expansion and contraction. The index is closely related to the "Activity Index" studied in S&W.[4] Our model based on this index postulates that changes in 12-month inflation, in addition to recent changes in inflation, also depend on current and past values of the CFNAI. That is, for $J = 12, 24$,

$$5) \quad \pi^{12}_{t+J} - \pi^{12}_t = \alpha + \beta(L)\,(\pi_t - \pi_{t-1}) + \gamma(L)a_t + \varepsilon_{t+J}\,,$$

where a_t denotes the value of the CFNAI at date t, and $\beta(L)$ and $\gamma(L)$ specify the number of lags in inflation and the index, respectively, included in the equation. We refer to this as model 2.

The remaining models we consider are based on the *diffusion index* methodology described in S&W. This methodology uses a small number of unobserved indexes that explain the movements in a large number of macroeconomic time series. Our implementation of the S&W methodology uses 154 data series, including data measuring production, labor market status, the strength of the household sector, inventories, sales, orders, financial market, money supply, and price data. The procedure that obtains the indexes processes the information in the 154 series so that each index is a weighted average of the series and each index is statistically independent of the others. We consider six indexes, $d_{1t}, d_{2t}, \ldots, d_{6t}$, which are ranked in descending order in terms of the amount of information embedded in them.

Our diffusion index models postulate that changes in 12-month inflation depend on recent changes in inflation, and current and past values of a number of diffusion indexes. That is, for $J = 12, 24$,

$$6) \quad \pi^{12}_{t+J} - \pi^{12}_t = \alpha + \beta(L)\,(\pi_t - \pi_{t-1}) + \sum_{i=1}^{K}\theta_i(L)d_{it} + \varepsilon_{t+J}\,,$$

where $K = 1, 2, \ldots, 6$, and $\beta(L)$ and $\theta_i(L)$ specify the number of lags in inflation and diffusion index i, respectively, included in the equation. As more indexes are included in the equation, more information about the 154 series is incorporated in the forecast. We refer to these as models 3, 4, ..., 8.

For all these models, we equate the forecasts of inflation with the conditional expectation implied by the model. That is, for $J = 12, 24$,

$$\hat{\pi}^{12}_{t+J} = E_t\pi^{12}_{t+J}\,.$$

We estimate all these models by ordinary least squares (OLS). In each case, we use the Bayes Information Criterion (BIC) to select the number of lags of inflation, the CFNAI, and the diffusion indexes. Intuitively, BIC selects the number of lags to improve the fit of the model without increasing by too much the sampling error in the lag coefficients. We allowed for the possibility that lags could vary from 0 to 11.

In real time, it is difficult to choose the appropriate model to use to form a forecast. To address this issue, we consider a forecasting model in which the forecast of inflation at any given date is the median of the forecasts of models 1 through 8 at that date.[5] This procedure has the advantage that it can be applied in real time. We call this the median model. Stock and Watson (2001) use a similar model. For convenience, we refer to the collection of models comprising models 1 through 8 plus the median model as Phillips curve models.

Model evaluation methodology

We evaluate the accuracy of the generalized Phillips curve models by comparing them with the naive model. We do this through various *simulated out-of-sample forecasting exercises*. These exercises involve constructing inflation forecasts that a model would have produced had it been used historically to generate forecasts of inflation. Two drawbacks of our approach, which also affect A&O and S&W, are that we do not use real-time data in our forecasts and we assume all the data are available up to the forecasting date. On a given date, particular data series may not yet be published. Also, many data series are revised after the initial release date. In our forecasting exercises, we calculate the CFNAI and diffusion indexes assuming all the series underlying the indexes are available up to the forecast date.[6] In practice, this is never the case and we must fill in missing data with estimates. Since we do not use real-time data to construct the CFNAI and diffusion indexes, we also abstract from problems associated with data revisions. We suspect that these drawbacks lead us to overstate the effectiveness of our CFNAI and diffusion index models. Data revision is also a problem for the lagged inflation and naive PCE models, since this price index is subject to revisions. It does not affect the CPI versions of these models, since the CPI is never revised.

To assess the accuracy of our various models, we first construct a measure of the average magnitude of the forecasting error. The measure we use is *root mean squared error* (RMSE). The RMSE for any forecast is the square root of the arithmetic average of the squared differences between the actual inflation rate and the predicted inflation rate over the period for which simulated forecasts are constructed. For $J = 12$, 24,

$$7)\quad RMSE = \left(\frac{1}{T} \sum_{t=1}^{T} \left[\pi_{t+J}^{12} - \hat{\pi}_{t+J}^{12} \right]^2 \right)^{1/2},$$

where T denotes the number of forecasts made over the period under consideration. We compare the forecast of a given Phillips curve model with that of the naive model by forming the ratio of the RMSE for the Phillips curve model to the RMSE for the naive model. We call this ratio the relative RMSE.

A ratio less than 1 thus indicates that the Phillips curve model is more accurate than the naive model. Subtracting 1 from the ratio and multiplying the result by 100 gives the percentage difference in RMSE between the two models. The RMSE might be strongly affected by one or two large outliers. We reworked our analysis using a measure of forecasting error that places equal weight on all forecasting errors and found that our findings are robust.[7] The RMSE statistics are subject to sampling variability and, consequently, are measured with error. In principle, we could use Monte Carlo methods to assess the magnitude of this error. However, this would require specifying an underlying data-generating process for all the variables in our analysis (more than 150 of them). One should keep this sampling error in mind when interpreting the results below.

The sample period of our analysis begins in 1967. We chose this date because it is the beginning date for the data used to construct the CFNAI and the diffusion indexes.[8] We estimate the forecasting equations using *rolling regressions*, a method that keeps the number of observations in the regression constant across forecasts. Since it excludes observations from the distant past, this approach can in principle accommodate the possibility of structural change in the data-generating process. We choose this sample length for the rolling regression procedure to be 15 years.[9]

Finally, we consider three distinct periods over which to evaluate the forecasts of the models: 1977–84, 1985–92, and 1993–2000. To compare our results with those in A&O, we also evaluate the overall performance of the models over the 1985–2000 period. To complete the analysis, we study the performance of the models over the entire 1977–2000 period as well. The 1977–84 period is one of high inflation volatility and general economic turbulence. The 1985–92 period is generally associated with a new monetary policy regime. This period also includes a mild recession. The 1993–2000 period witnessed uninterrupted economic expansion, stable monetary policy, and declining inflation.

Atkeson and Ohanian revisited

We estimated the Phillips curve models for the five sample periods and computed the associated RMSEs and the relative RMSEs. For models 1–8, we do not report all the results, just the results for the best models. We do this to demonstrate the potential forecasting capacity of these models. A&O report the performance of the best and worst models they look at across different lag lengths. S&W use BIC to select lag length and report the performance of all their models. All of these approaches suffer from the deficiency that in real time one may not know which is the best performing model. Our median model overcomes this deficiency. Table 1 displays the RMSE statistics of the best and median 12-month-ahead and 24-month-ahead forecasts for the five sample periods and four measures of inflation. The table also identifies the best performing models. The numbers in bold in the table indicate cases in which the naive model outperforms the Phillips curve models. Finally, for each case we report the RMSE for the naive model.

Regarding the 12-month-ahead forecasts in table 1, our findings are as follows. First, over the 1985–2000 period, essentially all the relative RMSEs are at least as large as 1. That is, the naive model outperforms all the Phillips curve models. This finding confirms the result reported in A&O that Phillips curve models have not performed well over the last 15 years. Second, while inflation forecasting appears to have been quite difficult over the last 15 years, for core PCE it has become a little easier in the most recent forecasting

TABLE 1

Forecasting the magnitude of inflation: Phillips curve models vs. naive model

	12 months ahead				24 months ahead			
Sample period	Naive model RMSE	Best performing model	Rel. RMSE	Median rel. RMSE	Naive model RMSE	Best performing model	Rel. RMSE	Median rel. RMSE
Core CPI								
1977:01–1984:12	2.360	2	0.768	0.885	3.802	4	0.930	0.868
1985:01–1992:12	0.667	1	0.985	**1.290**	0.780	1	0.894	**1.615**
1993:01–2000:12	0.341	**2**	**1.110**	**1.181**	0.705	2	0.765	0.768
1985:01–2000:12	0.530	**1**	**1.016**	**1.268**	0.744	1	0.903	**1.304**
1977:01–2000:12	1.430	2	0.891	0.927	2.278	**5**	**1.000**	0.906
Core PCE								
1977:01–1984:12	1.238	2	0.954	**1.033**	2.100	5	0.887	0.765
1985:01–1992:12	0.481	**1**	**1.409**	**1.412**	0.617	**1**	**1.221**	**1.197**
1993:01–2000:12	0.514	4	0.750	0.749	0.802	4	0.532	0.542
1985:01–2000:12	0.498	**1**	**1.188**	**1.109**	0.716	6	0.933	0.847
1977:01–2000:12	0.822	**2**	**1.048**	**1.052**	1.346	5	0.902	0.781
Total CPI								
1977:01–1984:12	2.674	2	0.687	0.765	4.525	4	0.744	0.696
1985:01–1992:12	1.489	1	0.982	0.982	1.695	1	0.981	**1.245**
1993:01–2000:12	0.716	**1**	**1.085**	**1.193**	1.032	**1**	**1.035**	**1.002**
1985:01–2000:12	1.168	**1**	**1.002**	**1.025**	1.403	1	0.996	**1.184**
1977:01–2000:12	1.815	2	0.865	0.845	2.853	6	0.954	0.795
Total PCE								
1977:01–1984:12	1.705	2	0.841	0.953	2.977	6	0.751	0.686
1985:01–1992:12	1.025	1	0.978	**1.012**	1.102	**1**	**1.029**	**1.279**
1993:01–2000:12	0.633	4	0.953	0.960	0.924	6	0.773	0.772
1985:01–2000:12	0.852	**1**	**1.003**	0.998	1.017	**1**	**1.020**	**1.098**
1977:01–2000:12	1.205	2	0.974	0.968	1.909	6	0.909	0.781

Notes: Fifteen-year rolling regression. RMSE is root mean squared error. Numbers in bold indicate cases in which the naive model outperforms the Phillips curve models.

period. In particular, the forecast by the best model and the median forecast have RMSEs 25 percent lower than the naive model over the 1993–2000 period. Note, however, that this pattern is not true for core CPI. Third, the Phillips curve models are generally better than the naive model in the 1977–84 period. This result is uniform across inflation measures, except for the median forecast for core PCE. In some cases the improvement is quite dramatic. For example, the best CFNAI model is more than 30 percent better than the naive model when forecasting total CPI. Even the median forecast is about 24 percent better than the naive model.

The results for the 24-month-ahead forecasts in table 1 suggest that, over longer horizons than 12 months, the Phillips curve models may more consistently outperform the naive model. In particular, the best models at forecasting core inflation outperform the naive model in the 1985–2000 period. However, the gains are not dramatic. For core CPI, the gain is roughly 10 percent, and for core PCE it is about 7 percent. The median forecasts for core CPI over this period fare worse, but they are better for core PCE with a gain of 15 percent over the naive model. In the recent 1993–2000 period, the gains over the naive model are more substantial. The best models improve relative to the naive model by 24 percent for core CPI and 47 percent for core PCE. We see similar gains for the median forecasts. Finally, there are across the board gains using Phillips curve models to forecast 24 months ahead for the 1977–84 period.

We can summarize these findings as follows. First, the naive model does poorly in the 1977–84 period and relatively well in the 1985–92 period, forecasting 12 months ahead. Second, the naive model does not do well forecasting PCE inflation in the recent 1993–2000 period. Finally, the naive model does better forecasting 12 months ahead than 24 months ahead.

We can attribute the first finding to an apparent structural change in the early 1980s and the consequent decline in inflation volatility in the post-1984 period compared with the previous period. This decline in volatility is evident in the pattern of RMSEs for the naive model in table 1 (also see figure 1).[10] Given that the naive model predicts no change in inflation, it should do better in a period of low inflation volatility than in a period of high volatility. It is unclear how the performance of the Phillips curve models is affected by inflation volatility. Nevertheless, we suspect that changes in inflation volatility are a contributing factor to the poor performance of the naive model in the 1977–84 period and its significant improvement in the most recent 15 years. Another factor that probably plays an important part in explaining our first finding is that forecasting models do relatively well in a stable environment. If the structure of the economy changes, then regression equations tend to forecast with more error. We suspect the change in structure in the early 1980s has a lot to do with a change in monetary policy regime around that time. We think volatility and structural stability change may explain the second finding as well. In particular, it appears that there was a further decline in core CPI volatility in the 1993–2000 period, which is not matched by core PCE.

We think one possible explanation of the improved performance of the Phillips curve models at the 24-month forecast horizon has to do with the sluggish response of the economy to monetary policy actions. It is generally understood that economic activity and inflation respond with a considerable lag to changes in monetary policy, and that inflation is more sluggish in its response than economic activity. If this is true then there may be less information about future inflation in the 12-month-ahead forecasts than in the 24-month-ahead forecasts. Note that as the forecast horizon is increased, forecasting performance in terms of RMSE generically worsens. We can see this by comparing the RMSEs of 12-month-ahead and 24-month-ahead forecasts of the naive model in table 1. Evidently, the forecast errors for the Phillips curve models deteriorate at a slower rate than the forecast errors for the naive model.

Forecasting direction

In the previous section we used the RMSE criterion to evaluate the models. This measure emphasizes the ability of a forecasting model to predict the magnitude of inflation. In this section, we consider a complementary approach to evaluating forecasting models, which emphasizes the forecasted direction of change of future inflation.

What do the models we have described have to say about direction of change of inflation? First, consider the naive model. Strictly speaking, according to equations 2 and 3, this model always predicts no change in inflation. In principle the martingale hypothesis, equation 1, on which the naive model is based, could be used to make forecasts about direction. Given the conditional distribution of inflation 12 months and 24 months ahead, we could assess the probability of an increase or decrease in inflation over these horizons and use this to make predictions about the direction of change. If this distribution is symmetric around the conditional mean, then the martingale hypothesis would suggest that the likelihood of an increase in inflation is always 50 percent. If the distribution is skewed, the odds of inflation changing in a particular direction would be better than a coin flip. The martingale hypothesis does not provide any information about the nature of the conditional distribution.

Deriving predictions about the direction of inflation changes from a Phillips curve model is more straightforward. The main difference from the naive model is that the conditional expectation of inflation 12 months and 24 months ahead is not constrained to equal current inflation. Consequently, we can infer the direction of change by making minimal assumptions about the distribution of the error terms in equations 4–6. Specifically, if these distributions are symmetric, then the direction of change is given by the sign of the difference between the conditional forecast and the current value of inflation.

Now we analyze the ability of our models to forecast direction. We assume the forecast errors are symmetrically distributed. Therefore, the naive model predicts inflation increases with probability 50 percent. We evaluate our Phillips curve models by assessing how well they can forecast direction relative to this baseline. Specifically, for a given Phillips curve model, let $\hat{D}^{12}_{t+J}$ be the predicted direction of change in inflation J periods ahead. We define $\hat{D}^{12}_{t+J}$ as follows for $J = 12, 24$:

$$8)\quad \hat{D}^{12}_{t+J} = \begin{cases} +1 & \text{if } E_t\pi^{12}_{t+J} > \pi^{12}_t \\ -1 & \text{otherwise} \end{cases},$$

where $\hat{D}^{12}_{t+J} = +1$ indicates a forecasted increase in inflation and $\hat{D}^{12}_{t+J} = -1$ indicates a decrease. Actual changes in inflation are defined analogously. Let D^{12}_{t+J} be the actual direction of change in inflation J periods ahead, for $J = 12, 24$,

$$D^{12}_{t+J} = \begin{cases} +1 & \text{if } \pi^{12}_{t+J} > \pi^{12}_t \\ -1 & \text{otherwise} \end{cases}.$$

We measure the directional change performance of a model by measuring the percentage of the directional

TABLE 2

Forecasting the direction of inflation changes

Sample period	12 months ahead: Best performing model	PDPC	Median PDPC	24 months ahead: Best performing model	PDPC	Median PDPC
Core CPI						
1977:01–1984:12	3	75.0	71.9	3	91.7	82.3
1985:01–1992:12	7	62.5	59.4	6	66.7	63.5
1993:01–2000:12	4	78.1	80.2	1	85.4	78.1
1985:01–2000:12	7	70.3	69.8	1	74.5	70.8
1977:01–2000:12	7	69.1	70.5	3	75.0	74.7
Core PCE						
1977:01–1984:12	2	79.2	69.8	2	90.6	87.5
1985:01–1992:12	1	61.5	**42.7**	6	61.5	52.1
1993:01–2000:12	8	69.8	69.8	1	90.6	82.3
1985:01–2000:12	1	64.1	56.3	6	70.8	67.2
1977:01–2000:12	2	67.0	60.8	2	72.2	74.0
Total CPI						
1977:01–1984:12	2	86.5	71.9	4	92.7	89.6
1985:01–1992:12	8	60.4	58.3	6	76.0	72.9
1993:01–2000:12	4	60.4	57.3	3	77.1	74.0
1985:01–2000:12	5	59.4	57.8	5	73.4	73.4
1977:01–2000:12	2	62.8	62.5	4	78.5	78.8
Total PCE						
1977:01–1984:12	3	89.6	77.1	4	94.8	93.8
1985:01–1992:12	1	56.3	52.1	4	68.8	62.5
1993:01–2000:12	5	71.9	67.7	5	80.2	76.0
1985:01–2000:12	7	62.0	59.9	5	74.0	69.3
1977:01–2000:12	7	65.6	65.6	5	80.2	77.4

Notes: Fifteen-year rolling regression. RMSE is root mean squared error. PDPC indicates percentage of directional predictions that are correct. Numbers in bold indicate failure with respect to the naive model.

predictions that are correct (PDPC) in a particular sample period. This percentage is defined as (for $J = 12, 24$),

$$PDPC = \frac{1}{T}\sum_{t=1}^{T} I\left\{\hat{D}^{12}_{t+J} = D^{12}_{t+J}\right\},$$

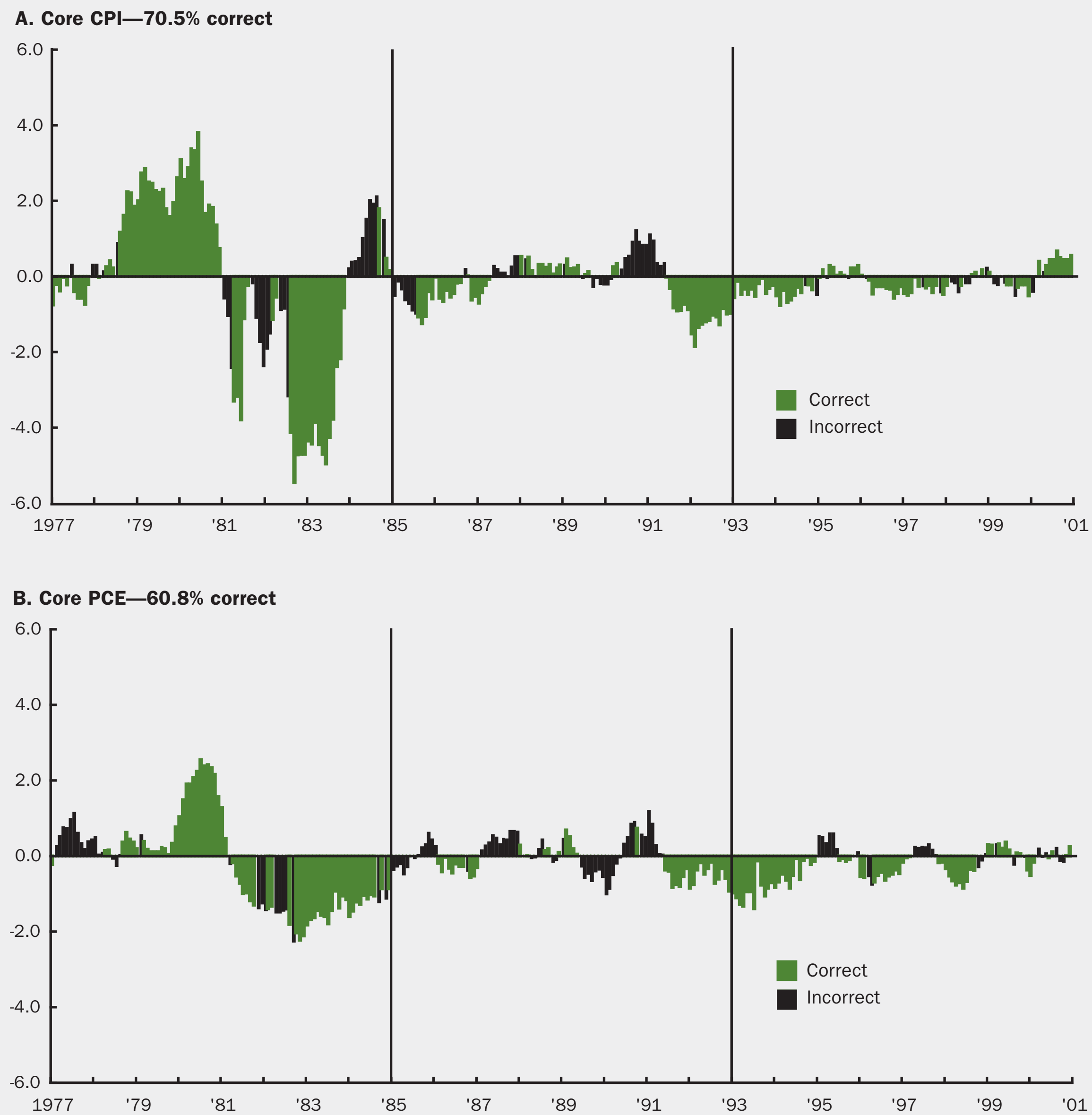
where I takes the value 1 when its argument is true (that is, $\hat{D}^{12}_{t+J} = D^{12}_{t+J}$), and 0 otherwise.

We used our estimates of the Phillips curve models computed for the RMSE comparisons in the previous section to make predictions about the direction of change of inflation according to equation 8. We report the findings for the best Phillips curve models and for the median model. Table 2 displays the 12- and 24-month-ahead directional predictions for the five sample periods and four measures of inflation. The table also identifies the best performing models. Numbers in bold indicate failure with respect to the naive model.

Our findings can be summarized as follows. It is immediately clear from the tables that the Phillips curve models predict direction in excess of 50 percent of the time for both 12-month and 24-month horizons in all but one case. Similar to their performance in terms of RMSE, these models are typically best at predicting directional change during the 1977–84 period and worst in the 1985–92 period. Interestingly, the best models at predicting directional changes are not the same as the best models in terms of RMSE. For example, model 4 (the model that includes d_{1t} and d_{2t}) provides the best 12-month-ahead forecasts of directional changes of core CPI over the 1993–2000 period. In terms of RMSE, model 2 provides the best forecasts over this sample period. Moreover, it is possible for a model to do well on directional changes while underperforming the naive model in terms of magnitude. In the example just given, the best directional change model is correct more than 78 percent of the time, but the best RMSE models in the corresponding period

FIGURE 2

Median model directional predictions of 12-month changes in 12-month core inflation

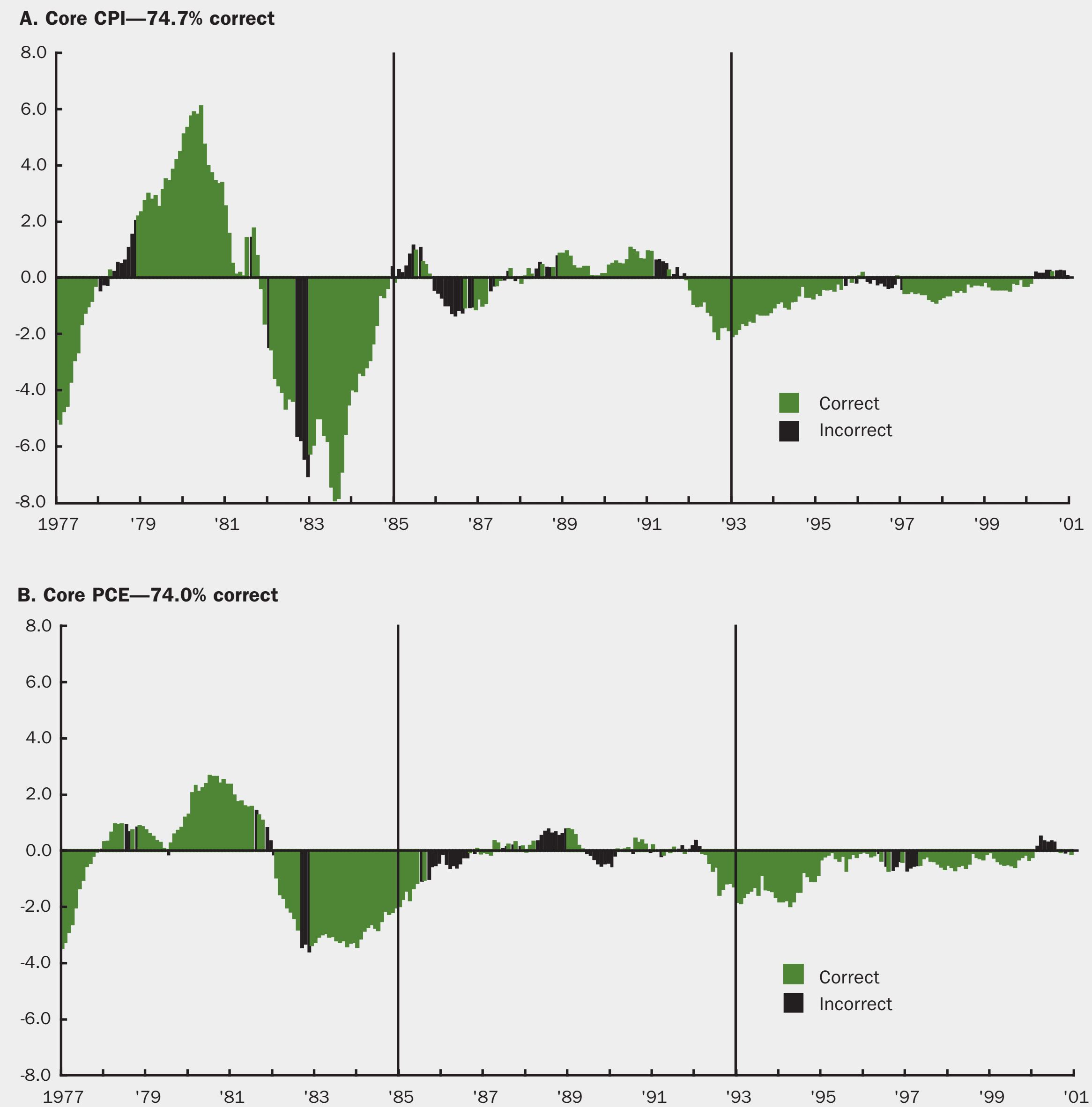


are worse than the naive model. Finally, the 24-month-ahead directional change forecasts perform better than the 12-month ahead forecasts.

Figures 2 and 3 provide information on when our directional change forecasts are correct. The bars indicate actual changes in core CPI and PCE inflation and the green bars indicate the correct directional predictions of the median model over the 1977–2000 period. The main lesson from these figures is that much of the success of the directional forecasts derives from periods in which there is a consistent trend in one direction or the other—the longer periods of consecutive increasing or decreasing inflation are associated with better directional forecasting. The relatively poor performance in the 1985–92 period may be partially due to the absence of a trend. As with our interpretation of the RMSE findings, we believe the change in monetary policy regime may also play a role. Interestingly, in the recent 1993–2000 period, despite the general downward trend in core CPI inflation, the one-year directional forecasts are able to correctly anticipate the brief episodes of increasing inflation.

FIGURE 3

Median model directional predictions of 24-month changes in 12-month core inflation



Can we improve on the naive model in difficult times?

Confirming the A&O findings, we show that the naive model has done quite well over the last 15 years in forecasting the magnitude of inflation. Over the same period, the Phillips curve models seem to provide information on the direction of changes in inflation. A natural question is whether we can combine these models to get a better forecast for magnitude. Intuitively, we should be able to do this by shaving the naive model forecasts up or down according to the directional predictions. In this section, we explore this idea and show that, indeed, it is possible to do somewhat better than the naive model.

We modify the naive model by adjusting its forecast in the direction predicted by a given Phillips curve or median model. That is, for $J = 12, 24$,

$$\hat{\pi}^{12}_{t+J} = \pi^{12}_{t} + \hat{D}^{12}_{t+J} \times v_t,$$

where $\hat{D}^{12}_{t+J}$ is defined in equation 8 and v_t is the adjustment factor. The intuition is that, for small enough

v_t, we should be able to improve on the naive model. In addition, we believe the magnitude of v_t should be related to recent changes in inflation. Consequently, we adjust the naive model by a percentage of the average inflation change in the recent past. That is, we assume v_t evolves as follows:

$$v_t = \lambda \times \sum_{j=t-N}^{t} \left| \pi_j^{12} - \pi_{j-12}^{12} \right| .$$

There is nothing in this approach that pins down v_t, and one may define v_t in other ways, provided that it is not too large. This formulation assumes symmetry in magnitude of increases and decreases in inflation. Choices of λ and N reflect the forecaster's belief in the relevance of past volatility of inflation for future volatility. For fixed N, intuition suggests that, for small enough λ, there will be at least a slight improvement over the naive model. We choose $\lambda = 0.1$ and N to correspond to the beginning of the regression sample. We call this the combination model.

Table 3, constructed in the same way as table 1, shows how well the combination model performs relative to the naive model. These results confirm our belief that we can improve on the naive model almost uniformly. For example, over the 1985–2000 period, the improvement of the best performing combination model for core CPI is about 7 percent and that for core PCE is about 3 percent for the 12-month horizon. For the 24-month horizon, the gains are 9 percent and 6 percent, respectively. Admittedly, these are not large improvements. The results for the median-based combination forecasts are less encouraging. The bad performance in the 1985–92 period seems to be driven by the relatively poor performance of the directional forecasts for both one-year and two-year horizons. The performance of the combination model may improve slightly by increasing λ by a small amount, but not by much.

TABLE 3

Forecasting the magnitude of inflation: Combination models vs. naive model

	12 months ahead				24 months ahead			
Sample period	Naive model RMSE	Best performing model	Rel. RMSE	Median rel. RMSE	Naive model RMSE	Best performing model	Rel. RMSE	Median rel. RMSE
Core CPI								
1977:01–1984:12	2.360	2	0.959	0.968	3.802	3	0.950	0.963
1985:01–1992:12	0.667	7	0.955	0.981	0.780	6	0.965	0.994
1993:01–2000:12	0.341	7	0.855	0.859	0.705	1	0.833	0.854
1985:01–2000:12	0.530	7	0.935	0.957	0.744	1	0.910	0.934
1977:01–2000:12	1.430	2	0.965	0.967	2.278	3	0.950	0.961
Core PCE								
1977:01–1984:12	1.238	5	0.954	0.962	2.100	3	0.945	0.950
1985:01–1992:12	0.481	1	0.992	**1.069**	0.617	6	0.970	**1.016**
1993:01–2000:12	0.514	1	0.944	0.942	0.802	1	0.910	0.925
1985:01–2000:12	0.498	1	0.967	**1.003**	0.716	6	0.943	0.960
1977:01–2000:12	0.822	7	0.966	0.972	1.346	3	0.948	0.952
Total CPI								
1977:01–1984:12	2.674	3	0.951	0.959	4.525	4	0.949	0.956
1985:01–1992:12	1.489	8	0.983	0.990	1.695	6	0.952	0.969
1993:01–2000:12	0.716	5	0.990	0.997	1.032	5	0.915	0.938
1985:01–2000:12	1.168	6	0.987	0.992	1.403	6	0.950	0.961
1977:01–2000:12	1.815	2	0.965	0.968	2.853	6	0.952	0.957
Total PCE								
1977:01–1984:12	1.705	3	0.939	0.958	2.977	5	0.941	0.946
1985:01–1992:12	1.025	1	0.988	**1.003**	1.102	5	0.934	0.961
1993:01–2000:12	0.633	5	0.944	0.962	0.924	5	0.911	0.925
1985:01–2000:12	0.852	7	0.982	0.992	1.017	5	0.924	0.946
1977:01–2000:12	1.205	2	0.961	0.970	1.909	5	0.938	0.946

Notes: Fifteen-year rolling regression. RMSE is root mean squared error. Numbers in bold indicate cases in which the naive model outperforms the combination model.

Conclusion

We can summarize our main results as follows. First, we show that the A&O findings hold for a broader class of models than they studied, as well as for a longer forecasting horizon. However, they do not hold for the 1977–84 period. We extend their analysis to core PCE and show that the naive model does better over the sample period considered by A&O at the one-year horizon, but not at the two-year horizon. In the 1993–2000 period, the Phillips curve models perform well at forecasting core PCE for both horizons. Second, we show that Phillips curve models have predictive power for the direction of change in inflation. This is particularly true in the 1977–84 and 1993–2000 periods. However, in the 1985–92 period, the gain over the naive model is quite modest. Third, in most cases it is possible to combine the information in the directional forecasts with the naive model to improve on the latter model's forecasts.

A common thread in our results is the relatively poor performance of the Phillips curve models in the middle period, in terms of both magnitude and direction. We believe this is due to a reduction in inflation volatility and the change in monetary policy operating characteristics that took effect in this time.

Our findings suggest the following policy recommendation. If we expect the current monetary "regime" to persist, then we can have some degree of confidence in the Phillips curve models going forward. On the other hand, if we suspect that a regime shift has recently occurred, then we should be skeptical of the Phillips curve forecasts. In any case, there may be some directional information in these forecasts, and we can use this to improve on naive forecasts.

Our findings suggest that more empirical and theoretical work is necessary to come to a complete answer to the question raised in the title to this article. An equivalent way of posing this question is to ask: Why does inflation behave like a martingale over some periods while at other times it does not? We have suggested some possible explanations. Empirically, we need to assess the robustness of our results to cross-country analysis. For example, here we have only one regime change and, hence, only one observation for the regime-switch hypothesis. Ultimately, assessing the plausibility of various possible explanations will require developing a fully specified theoretical model. Such work may shed light on the connection between monetary policy and aggregate outcomes, as well as the nature of the price-setting mechanism.

NOTES

[1]For a recent discussion of the intellectual history of the Phillips curve and NAIRU, see Gordon (1997).

[2]See Bernanke and Mihov (1998), Bordo and Schwartz (1997), and Strongin (1995) for discussions of monetary regimes. These papers argue that during the Volker chairmanship of the Board of Governors from 1979–87 monetary policy shifted, in terms of operating procedures and the Fed's increased willingness to combat inflation. Furthermore, Bernanke and Mihov (1998) estimate monetary policy rules and can reject the hypothesis of parameter stability for dates in the 1980s.

[3]Suppose there are K lags in the equation, then $\beta(L)x_t = \beta_0 x_t + \beta_1 x_{t-1} + \ldots \beta_K x_{t-K}$, where the β parameters are scalars.

[4]For more details on the CFNAI, see www.chicagofed.org/economicresearchanddata/national/index.cfm.

[5]With eight models, the median is the average forecast of the fourth and fifth ranked forecasts.

[6]One implication of this procedure is that the historical path of the indexes may change between forecast dates.

[7]The alternative measure we used was mean absolute value error. For $J = 12, 24$, this is expressed as $(1/T)\sum_{i=1}^{T}\left|\pi_{i+J}^{12} - \hat{\pi}_{i+J}^{12}\right|$.

[8]A&O use several sample periods for their analysis. When they consider unemployment-rate-based Phillips curves, their sample begins in 1959. When they consider CFNAI-based Phillips curves, their sample begins in 1967. S&W begin their analysis in 1959.

[9]We also considered estimating the forecasting equations using all the data from 1967 up to the forecast date. The results obtained indicated either very similar forecast performance or, in a few cases, a slight deterioration relative to the rolling regression procedure.

[10]S&W report evidence supporting this hypothesis. They find that unemployment-rate-based Phillips curve forecasting models exhibit parameter instability during the 1980s.

REFERENCES

Atkeson, Andrew, and Lee E. Ohanian, 2001, "Are Phillips curves useful for forecasting inflation?," Federal Reserve Bank of Minneapolis, *Quarterly Review*, Vol. 25, No. 1, Winter, pp. 2–11.

Bernanke, Ben, and Ilian Mihov, 1998, "Measuring monetary policy," *Quarterly Journal of Economics*, Vol. 113, No. 3, August, pp. 869–902.

Bordo, Michael, and Anna Schwartz, 1997, "Monetary policy regimes and economic performance: The historical record," National Bureau of Economic Research, working paper, No. 6201.

Gordon, Robert, 1997, "The time-varying NAIRU and its implications for economic policy," *Journal of Economic Perspectives*, Vol. 11, No. 1, Winter, pp. 11–32.

Stock, James H., and Mark W. Watson, 2001, "Forecasting output and inflation: The role of asset prices," Princeton University, manuscript, February.

__________, 1999, "Forecasting inflation," *Journal of Monetary Economics*, Vol. 44, pp. 293–335.

Strongin, Steven, 1995, "The identification of monetary policy disturbances: Explaining the liquidity puzzle," *Journal of Monetary Economics*, Vol. 35, pp. 463–497.