3

Forecasting Inflation

Phillips Curve Inflation Forecasts

James H. Stock and Mark W. Watson

1. Introduction

Inflation is hard to forecast. There is now considerable evidence that Phillips curve forecasts do not improve upon good univariate benchmark models. Yet the backward-looking Phillips curve remains a workhorse of many macroeconomic forecasting models and continues to be the best way to understand policy discussions about the rates of unemployment and inflation.

After some preliminaries set forth in section 2, this paper begins its analysis in section 3 by surveying the past fifteen years of literature (since 1993) on inflation forecasting, focusing on papers that conduct a pseudo out-of-sample forecast evaluation.[1] A milestone in this literature is Atkeson and Ohanian (2001), who considered a number of standard Phillips curve forecasting models and showed that none improve upon a four-quarter random walk benchmark over the period 1984–1999. As we observe in this survey, Atkeson and Ohanian deserve the credit for forcefully making this point; however, their finding has precursors dating back at least to 1994. The literature after Atkeson and Ohanian finds that their specific result depends rather delicately on the sample period and the forecast horizon. If, however, one uses other univariate benchmarks (in particular, the unobserved components-stochastic volatility model of Stock and Watson (2007)), the broader point of Atkeson and Ohanian—that, at least since 1985, Phillips curve forecasts do not outperform univariate benchmarks on average—has been confirmed by several studies. The development of this literature is illustrated empirically using six prototype inflation forecasting models: three univariate models, two backward-looking Phillips curve models—Gordon's (1990) "triangle" model and an autoregressive-distributed lag model using the unemployment rate—and a model using the term spread, specifically the yield spread between one-year Treasury bonds and 90-day Treasury bills.

It is difficult to make comparisons across papers in this literature because the papers use different sample periods, different inflation series, and different benchmark models, and the quantitative results in the literature are curiously dependent upon these details. In section 4, we therefore undertake an empirical study that aims to unify and to assess the results in the literature using quarterly U.S. data from 1953:Q1–2008:Q1. This study examines the pseudo out-of-sample performance of a total of 192 forecasting procedures (157 distinct models and 35 combination forecasts), including the six prototype models of section 3, applied to forecasting five different inflation measures (CPI-all, CPI-core, PCE-all, PCE-core, and the GDP deflator). This study confirms the main qualitative results of the literature, although some specific results are found not to be robust. Our study also suggests an interpretation of why the literature's conclusions strongly depend on the sample period. Specifically, one of our key findings is that the performance of Phillips curve forecasts is episodic: there are times, such as the late 1990s, when Phillips curve forecasts improved upon using univariate forecasts, but there are other times (such as the mid-1990s) when a forecaster would have been better off using a univariate forecast. This finding provides a rather more nuanced interpretation of Atkeson and Ohanian's (2001) conclusion concerning Phillips curve forecasts, one that is consistent with the sensitivity of findings in the literature to the sample period.

A question that is both difficult and important is what this episodic performance implies for an inflation forecaster today. On average, over the past 15 years, it has been very hard to beat the best univariate model using any multivariate inflation forecasting model (Phillips curve or otherwise). But suppose you are told that next quarter the economy would plunge into recession, with the unemployment rate jumping by 2 percentage points. Would you change your inflation forecast? The literature is now full of formal statistical evidence suggesting that this information should be ignored, but we suspect that an applied forecaster would nevertheless revise downward his or her forecast of inflation over the one- to two-year horizon. In the final section, we suggest some reasons why this revision might be justified.

2. Notation, Terminology, Families of Models, and Data

This section provides preliminary details concerning the empirical analysis and gives the six prototype inflation forecasting models that will be used in section 3 as a guide to the literature. We begin by reviewing some forecasting terminology.

Terminology

h-period inflation. Inflation forecasting tends to focus on the one-year or two-year horizons. We denote h-period inflation by $\pi_t^h = h^{-1}\sum_{i=0}^{h-1}\pi_{t-i}$, where π_t is the quarterly rate of inflation at an annual rate; that is, $\pi_t = 400\ln(P_t/P_{t-1})$ (using the log approximation), where P_t is the price index in quarter t. Four-quarter inflation at date t is $\pi_t^4 = 100\ln(P_t/P_{t-4})$, the log approximation to the percentage growth in prices over the previous four quarters.

Direct and iterated forecasts. There are two ways to make an h-period ahead model-based forecast. A direct forecast has π_{t+h}^h as the dependent variable and t-dated variables (variables observed at date t) as regressors; for example, π_{t+h}^h could be regressed on π_t^h and the date-t unemployment rate (u_t). At the end of the sample (date T), the forecast of π_{T+h}^h is computed "directly" using the estimated forecasting equation. In contrast, an iterated forecast is based on a one-step ahead model; for example, π_{t+1} could be regressed on π_t, which is then iterated forward to compute future conditional means of π_s, $s > T + 1$, given data through time t. If predictors other than past π_t are used, then this requires constructing a subsidiary model for the predictor, or alternatively, modeling π_t and the predictor jointly—for example, as a vector autoregression (VAR)—and iterating the joint model forward.

Pseudo out-of-sample forecasts; rolling and recursive estimation. Pseudo out-of-sample forecasting simulates the experience of a real-time forecaster by performing all model specification and estimation using data through date t, making a h-step ahead forecast for date $t + h$, then moving forward to date $t + 1$ and repeating this through the sample.[2]

Pseudo out-of-sample forecast evaluation captures model specification uncertainty, model instability, and estimation uncertainty, in addition to the usual uncertainty of future events.

Model estimation can either be rolling (using a moving data window of fixed size) or recursive (using an increasing data window, always starting with the same observation). In this paper, rolling estimation is based on a window of ten years, and recursive estimation starts in 1953:Q1 or, for series starting after 1953:Q1, the earliest possible quarter.

Root mean squared error and rolling RMSE. The root mean squared forecast error (RMSE) of h-period ahead forecasts made over the period t_1 to t_2 is

$$(1)\qquad RMSE_{t_1,t_2} = \sqrt{\frac{1}{t_2 - t_1 + 1}\sum_{t=t_1}^{t_2}\left(\pi^h_{t+h} - \pi^h_{t+h|t}\right)^2},$$

where $\pi^h_{t+h|t}$ is the pseudo out-of-sample forecast of π^h_{t+h} made using data through date t. This paper uses rolling estimates of the RMSE, which are computed using a weighted centered 15-quarter window:

$$(2)\qquad rolling\,RMSE(t) = \sqrt{\sum_{s=t-7}^{t+7} K(|s-t|/8)\left(\pi^h_{s+h} - \pi^h_{s+h|s}\right)^2 \Big/ \sum_{s=t-7}^{t+7} K(|s-t|/8)},$$

where K is the biweight kernel, $K(x) = (15/16)(1 - x^2)^2\mathbf{1}(|x| \leq 1)$.

Prototypical Inflation Forecasting Models

Single-equation inflation forecasting models can be grouped into four families: (1) forecasts based solely on past inflation; (2) forecasts based on activity measures ("Phillips curve forecasts"); (3) forecasts based on the forecasts of others; and (4) forecasts based on other predictors. This section lays out these families and provides prototype examples of each.

(1) ***Forecasts based on past inflation.*** This family includes univariate time series models such as autoregressive integrated moving average (ARIMA) models and nonlinear or time-varying univariate models. We also include in this family of forecasts those in which one or more inflation measure, other than the series being forecasted, is used as a predictor; for example, past Consumer Price Index (CPI) core inflation or past growth in wages could be used to forecast CPI-all inflation.

Three of our prototype models come from this family and serve as forecasting benchmarks. The first is a direct autoregressive (AR) forecast, computed using the direct autoregressive model,

$$\text{(3)}\quad \pi^h_{t+h} - \pi_t = \mu^h + \alpha^h(\mathrm{L})\Delta\pi_t + v^h_{t+h}, \qquad \text{(AR(AIC))}$$

where μ^h is a constant, $\alpha^h(\mathrm{L})$ is a lag polynomial written in terms of the lag operator L, v^h_{t+h} is the h-step ahead error term (we will use v generically to denote regression error terms), and the superscript h denotes the quantity for the h-step ahead direct regression. In this prototype AR model, the lag length is determined by the Akaike Information Criterion (AIC) over the range of 1 to 6 lags. This specification imposes a unit autoregressive root.

The second prototype model is the Atkeson-Ohanian (2001) random walk model, in which the forecast of the four-quarter rate of inflation, π^4_{t+4}, is the average rate of inflation over the previous four quarters, π^4_t (Atkeson and Ohanian only considered four-quarter ahead forecasting). The Atkeson-Ohanian model thus is,

$$\text{(4)}\quad \pi^4_{t+4} = \pi^4_t + v^4_{t+4} \qquad (AO).$$

The third prototype model is the Stock-Watson (2007) unobserved components-stochastic volatility (UC-SV) model, in which π_t has a stochastic trend τ_t, a serially uncorrelated disturbance η_t, and stochastic volatility:

$$\text{(5)}\quad \pi_t = \tau_t + \eta_t, \text{ where } \eta_t = \sigma_{\eta,t}\zeta_{\eta,t} \qquad \text{(UC-SV)}$$

$$\text{(6)}\quad \tau_t = \tau_{t-1} + \varepsilon_t, \text{ where } \varepsilon_t = \sigma_{\varepsilon,t}\zeta_{\varepsilon,t}$$

$$\text{(7)}\quad \ln \sigma^2_{\eta,t} = \ln \sigma^2_{\eta,t-1} + \nu_{\eta,t}$$

$$\text{(8)}\quad \ln \sigma^2_{\varepsilon,t} = \ln \sigma^2_{\varepsilon,t-1} + \nu_{\varepsilon,t}$$

where $\zeta_t = (\zeta_{\eta,t}, \zeta_{\varepsilon,t})$ is independent and identically distributed (i.i.d.). $\mathrm{N}(0, \mathrm{I}_2)$, $\nu_t = (\nu_{\eta,t}, \nu_{\varepsilon,t})$ is i.i.d. $\mathrm{N}(0, \gamma\mathrm{I}_2)$, ζ_t, and ν_t are independently distributed, and γ is a scalar parameter. Although η_t and ε_t are conditionally normal given $\sigma_{\eta,t}$ and $\sigma_{\varepsilon,t}$, unconditionally these are random mixtures of normal random variables and can have heavy tails. This is a one-step ahead model and forecasts are iterated. The UC-SV model has only one parameter, γ, which controls the smoothness of the stochastic volatility process. Throughout, we follow Stock and Watson (2007) and set $\gamma = 0.04$.

(2) ***Phillips curve forecasts.*** We interpret Phillips curve forecasts broadly to include forecasts produced using an activity variable, such as the unemployment rate, an output gap, or output growth, perhaps in conjunction with other variables, to forecast inflation or the change in inflation. This family includes both backward-looking Phillips curves and New Keynesian Phillips curves, although the latter appear infrequently (and only recently) in the inflation forecasting literature.

We consider two prototype Phillips curve forecasts. The first is Gordon's (1990) "triangle model," which in turn is essentially the model in Gordon (1982) with minor modifications.[3] In the triangle model, inflation depends on lagged inflation, the unemployment rate u_t, and supply shock variables z_t:

$$\pi_{t+1} = \mu + \alpha^G(L)\pi_t + \beta(L)u_{t+1} + \gamma(L)z_t + v_{t+1}. \qquad \text{(triangle)} \tag{9}$$

The prototype triangle model used here is that in Gordon (1990), in which (9) is specified using the contemporaneous value plus 4 lags of u_t (total civilian unemployment rate ages 16+ years, seasonally adjusted), contemporaneous value plus 4 lags of the rate of inflation of food and energy prices (computed as the difference between the inflation rates in the deflator for "all-items" personal consumption expenditure (PCE) and the deflator for PCE less food and energy), lags 1 through 4 of the relative price of imports (computed as the difference of the rates of inflation of the GDP deflator for imports and the overall GDP deflator), two dummy variables for the Nixon wage-price control period, and 24 lags of inflation, where $\alpha^G(L)$ imposes the step-function restriction that the coefficients are equal within the groups of lags 1–4, 5–8, …, 21–24, and also that the coefficients sum to one (a unit root is imposed).

Following Gordon (1998), forecasts based on the triangle model (9) are iterated using forecasted values of the predictors, where those forecasts are made using subsidiary univariate AR(8) models of u_t, food and energy inflation, and import inflation.

The second prototype Phillips curve model is direct version of (9) without the supply shock variables and without the step-function restriction on the coefficients. This model is an autoregressive distributed lag (ADL) model in which forecasts are computed using the direct regression,

$$\pi^h_{t+h} - \pi_t = \mu^h + \alpha^h(L)\Delta\pi_t + \beta^h(L)u_t + v^h_{t+h}, \qquad \text{(ADL-}u\text{)} \tag{10}$$

where $\alpha^h(L)$ and $\beta^h(L)$ are unrestricted with degrees chosen separately by AIC (maximum lag of 4), and (like the triangle model) the ADL-u specification imposes a unit root in the autoregressive dynamics for π_t.

(3) ***Forecasts based on forecasts of others.*** The third family computes inflation forecasts from explicit or implicit inflationary expectations or forecasts of others. These forecasts include regressions based on implicit expectations derived from asset prices, such as forecasts extracted from the term structure of nominal Treasury debt (which by the Fisher relation should embody future inflation expectations) and forecasts extracted from the Treasury Inflation-Protected Securities (TIPS) yield curve. This family also includes forecasts based on explicit forecasts of others, such as median forecasts from surveys such as the Survey of Professional Forecasters.

Our prototypical example of forecasts in this family is a modification of the Mishkin (1990) specification, in which the future change in inflation is predicted by a matched-maturity spread between the interest rates on comparable government debt instruments, with no lags of inflation. Here we consider direct 4-quarter ahead forecasts based on an ADL model using as a predictor the interest spread, $spread1\_90_t$, between one-year Treasury bonds and 90-day Treasury bills:

$$\pi^4_{t+4} - \pi_t = \mu + \alpha(L)\Delta\pi_t + \beta(L)spread1\_90_t + v^4_{t+4}. \qquad \text{(ADL-}spread\text{)} \tag{11}$$

We emphasize that Mishkin's (1990) regressions appropriately use term spread maturities matched to the change in inflation being forecasted, which for (11) would be the change in inflation over quarters $t + 2$ to $t + 4$, relative to $t + 1$. (A matched maturity alternative to $spread1\_90_t$ in (11) would be the spread between one-year Treasuries and the federal funds rate, however those instruments have different risks.) Because the focus of this paper is Phillips curve regressions we treat this regression simply as an example of this family and provide references to recent studies of this family in section 3.3.

(4) ***Forecasts based on other predictors.*** The fourth family consists of inflation forecasts that are based on variables other than activity or expectations variables. An example is a 1970s-vintage monetarist model in which M1 growth is used to forecast inflation. Forecasts in this fourth family perform sufficiently poorly relative to the three other approaches that these play negligible roles both in the literature and in current practice, so to avoid distraction we do not track a model in this family as a running example.

Data and transformations

The data set is quarterly for the United States from 1953:Q1–2008:Q1. Monthly data are converted to quarterly data by computing the average value for the three months in the quarter prior to any other transformations; for example, quarterly CPI is the average of the three monthly CPI values, and quarterly CPI inflation is the percentage growth (at an annual rate, using the log approximation) of this quarterly CPI.

We examine forecasts of five measures of price inflation: the GDP price deflator (PGDP), the CPI for all items (CPI-all), CPI excluding food and energy (CPI-core), the personal consumption expenditure deflator (PCE-all), and the personal consumption expenditure deflator excluding food and energy (PCE-core).

In addition to the six prototype models, in section 4 we consider forecasts made using a total of 15 predictors, most of which are activity variables (GDP, industrial production, housing starts, the capacity utilization rate, etc.). The full list of variables and transformations is given in the appendix.

Gap variables. Consistent with the pseudo out-of-sample forecasting philosophy, the activity gaps used in the forecasting models in this paper are all one-sided. Following Stock and Watson (2007), gaps are computed as the deviation of the series (for example, log GDP) from a symmetric two-sided moving average (MA(80)) approximation to the optimal lowpass filter with pass band corresponding to periodicities of at least 60 quarters. The one-sided gap at date t is computed by padding observations at dates $s > t$ and $s < 1$ with iterated forecasts and backcasts based on an AR(4), estimated recursively through date t.

3. An Illustrated Survey of the Literature on Phillips Curve Forecasts, 1993–2008

This section surveys the literature during the past fifteen years (since 1993) on inflation forecasting in the United States. The criterion for inclusion in this survey is providing empirical evidence on inflation forecasts (model- and/or survey-based) in the form of a true or pseudo out-of-sample forecast evaluation exercise. Such an evaluation can use rolling or recursive forecasting methods based on final data, it can use rolling or recursive methods using real-time data, or it can use forecasts actually produced and recorded in real time such as survey forecasts. Most of the papers discussed here focus on forecasting at horizons of policy relevance, one or two years. Primary interest is in forecasting overall consumer price inflation (PCE, CPI), core inflation, or economy-wide inflation (GDP deflator). There is little work on forecasting producer prices, although a few papers consider producer prices as a predictor of headline inflation.

This survey also discusses some papers in related literatures; however, we do not attempt a comprehensive review of those related literatures. One such literature concerns the large amount of interesting work that has been done on inflation forecasting in countries other than the United States: see Rünstler (2002), Hubrich (2005), Canova (2007), and Diron and Mojon (2008) for recent contributions and references. Another closely related literature concerns in-sample statistical characterizations of changes in the univariate and multivariate inflation process in the United States (e.g, Taylor 2000; Brainard and Perry 2000; Cogley and Sargent 2002, 2005; Levin and Piger 2004; and Pivetta and Reis 2007) and outside the United States (e.g., the papers associated with the European Central Bank Inflation Persistence Network 2007). There is in turn a literature that asks whether these changes in the inflation process can be attributed, in a quantitative (in-sample) way, to changes in monetary policy; papers in this vein include Estrella and Fuhrer (2003), Roberts (2004), Sims and Zha (2006), and Primiceri (2006). A major theme of this survey is time-variation in the Phillips curve from a forecasting perspective, most notably at the end of the disinflation of the early 1980s but more subtly throughout the post-1984 period. This time-variation is taken up in a great many papers; for example, papers estimating a time-varying NAIRU and time variation in the slope of the Phillips curve. In addition, there is a massive theoretical and empirical literature that develops and analyzes the New Keynesian Phillips curve (Roberts 1995). Papers in these literatures, however, are only discussed in passing unless they have a pseudo out-of-sample forecasting component.

The 1990s: Warning Signs

The great inflation and disinflation of the 1970s and the 1980s was the formative experience that dominated the minds and models of inflation forecasters through the 1980s and early 1990s, both because of the forecasting failures of 1960s-vintage ("non-accelerationist") Phillips curves and, more mechanically, because most of the variation in the data comes from that period. The dominance of this episode is evident in figure 3.1,

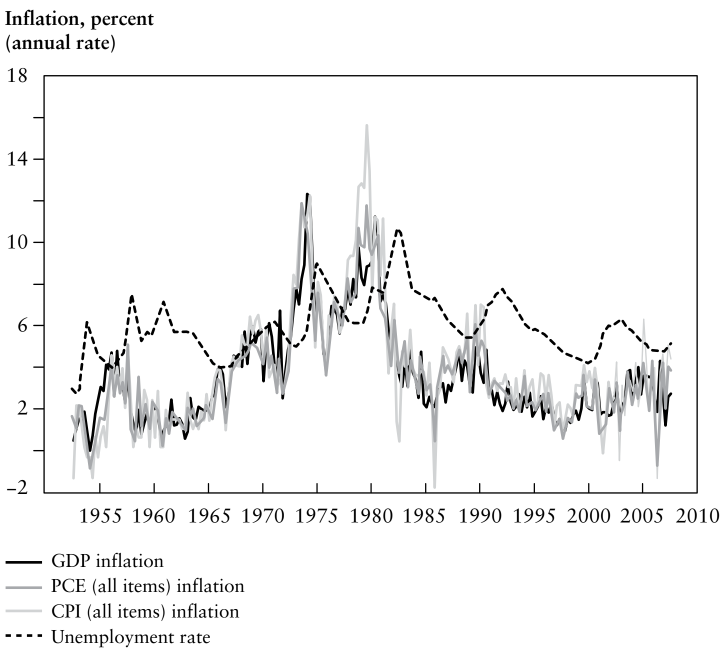


Figure 3.1
Quarterly U.S. Price Inflation at an Annual Rate as Measured by the GDP Deflator, PCE-All and CPI-All, and the Rate of Unemployment, 1953:Q1–2008:Q1

which plots the three measures of headline inflation (GDP, PCE-all, and CPI-all) from 1953:Q1 to 2007:Q4, along with the unemployment rate.

By the early 1980s, despite theoretical attacks on the backward-looking Phillips curve, Phillips curve forecasting specifications had coalesced around the Gordon (1982) triangle model (9) and variants. Figure 3.2 plots the rolling RMSE of the four-quarter ahead pseudo out-of-sample forecast of CPI-all inflation, computed using (2), for the recursively estimated AR(AIC) benchmark (3), the triangle model (9), and the ADL-*u* model (10). As can be seen in figure 3.2, these "accelerationist" Phillips curve specifications (unlike their non-accelerationist ancestors) did in fact outperform the AR(AIC) benchmark during the 1970s and 1980s.

While the greatest success of the triangle model and the ADL-*u* model was forecasting the fall in inflation during the early 1980s subsequent to the spike in the unemployment rate in 1980, in fact the triangle and ADL-*u* models improved upon the AR benchmark nearly uniformly from 1965 through 1990. The main exception occurred around 1986, when there was a temporary decline in oil prices. The four-quarter ahead pseudo out-of-sample forecasts produced by the AR(AIC), triangle, and

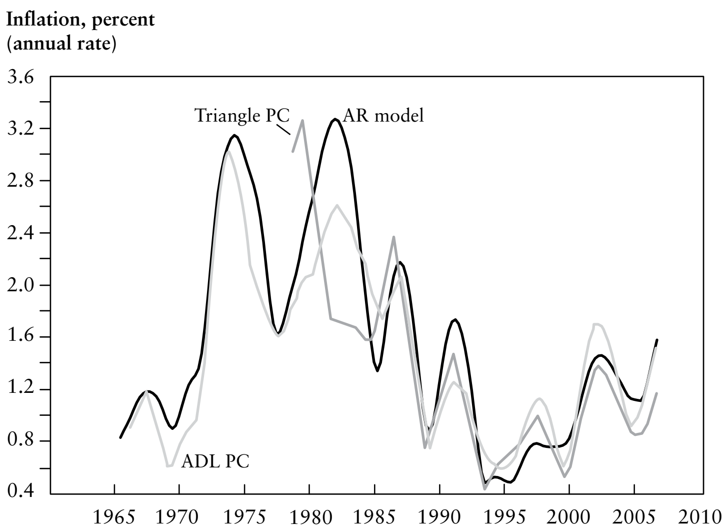


Figure 3.2
Rolling Root Mean Squared Errors for CPI-All Inflation Forecasts: AR(AIC), Triangle Model (constant NAIRU), and ADL-*u* Model

ADL-*u* models are shown respectively in panels (a)-(c) of figure 3.3. As can be seen in figure 3.3, the triangle model predicted too much too late: it initially failed to forecast the decline in inflation in 1986, then predicted inflation to fall further than it actually did. Interestingly, unlike the AR(AIC) and ADL-*u* models, triangle model forecasts did not over-extrapolate the decline in inflation in the early 1980s.

Stockton and Glassman (1987) documented the good performance of a triangle model based on the Gordon (1982) specification of the triangle model over the 1977–1984 period (they used the Council of Economic Advisors output gap instead of the unemployment rate and a 16-quarter, not 24-quarter, polynomial distributed lag). They reported a pseudo out-of-sample relative RMSE of the triangle model, relative to an AR(4) model of the change in inflation, of 0.80 (eight-quarter ahead iterated forecasts of inflation measured by the Gross Domestic Business Product fixed-weight deflator).[4] Notably, Stockton and Glassman (1987) also emphasized that there seem to be few good competitors to this model: a variety of monetarist models, including some that incorporate expectations of money growth, all performed worse—in some cases, much worse—than the AR(4) benchmark. This said, the gains from using a Phillips curve forecast over the second half of the 1980s were slimmer than during the 1970s and early 1980s.

The earliest documentation of this relative deterioration of Phillips curve forecasts of which we are aware is a little-known (two Google Scholar cites) working paper by Jaditz and Sayers (1994). They undertook a pseudo out-of-sample forecasting exercise of CPI-all inflation using industrial production growth, the PPI, and the 90-day Treasury Bill rate in a VAR and in a vector error correction model (VECM), with a forecast period of 1986-1991 and a forecast horizon of one month. They reported a relative RMSE of .985 for the VAR and a relative mean squared error (MSE) in excess of one for the VECM, relative to an AR(1) benchmark.

Cecchetti (1995) also provided early evidence of instability in Phillips curve forecasts. However, that instability was apparent only using in-sample break tests and did not come through in his pseudo out-of-sample forecasting evaluation because of his forecast sample period. Cecchetti considered forecasts of CPI-all at horizons of one–four years based on

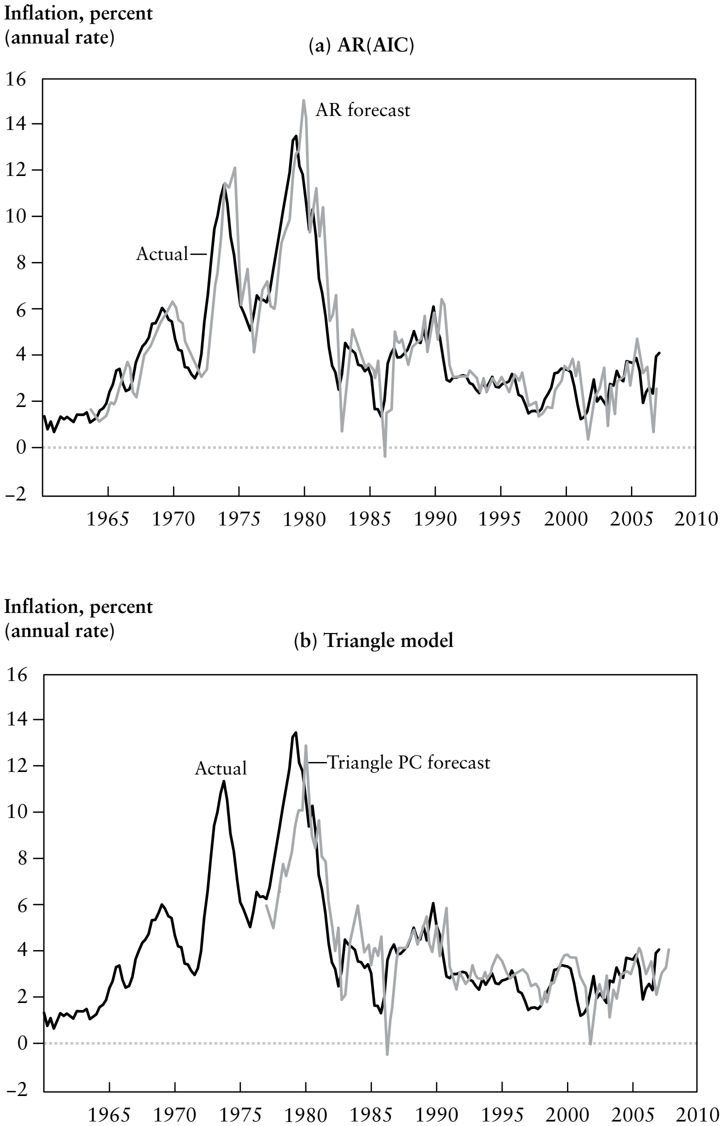


Figure 3.3
CPI-All Inflation and Pseudo Out-of-Sample Forecasts

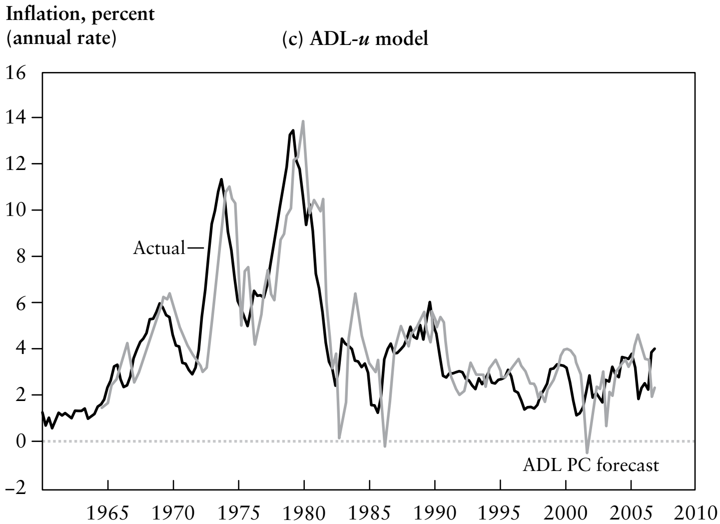


Figure 3.3 (continued)

18 predictors, entered separately, for two forecast periods, 1977–1994 ten-year rolling window) and 1987–1994 (five-year rolling window). Inspection of figure 3.2 indicates that Phillips curve forecasts did well on average over both of these samples, but that the 1987–1994 period was atypical of the post-1984 experience in that it is dominated by the relatively good performance of Phillips curve forecasts during the 1990 recession. Despite the good performance of Phillips curve forecasts over this period, using in-sample break tests Cecchetti (1995) found multiple breaks in the relationship between inflation and (separately) unemployment, the employment/population ratio, and the capacity utilization rate. He also found that good in-sample fit is essentially unrelated to subsequent pseudo out-of-sample forecasting performance.

Stock and Watson (1999) undertook a pseudo out-of-sample forecasting assessment of CPI-all and PCE-all forecasts at the one-year horizon using (separately) 168 economic indicators, of which 85 were measures of real economic activity (industrial production growth, unemployment, and so on). They considered recursive forecasts computed over two subsamples, 1970–1983 and 1984–1996. The split sample evidence indicated major changes in the relative performance of predictors in the two subsamples; for example, the RMSE of the forecast based on the unemployment rate, relative to the AR benchmark, was .89 in the 1970–1983 sample but 1.01 in the 1984–1996 sample. Using in-sample test statistics, they also found structural breaks in the inflation-unemployment relationship, although interestingly these breaks were more detectable in the coefficients on lagged inflation in the Phillips curve specifications than on the activity variables.

Cecchetti, Chu, and Steindel (2000) examined CPI inflation forecasts at the two-year horizon using (separately) 19 predictors, including activity indicators. They reported dynamic forecasts in which future values of the predictors are used to make multiperiod ahead forecasts (future employment is treated as known at the time the forecast is made, so these are not pseudo out-of-sample). Strikingly, they found that activity-based dynamic forecasts (unemployment, employment-population ratio, and capacity utilization rate) typically underperformed the AR benchmark over this period at the one-year horizon.

Brayton, Roberts, and Williams (1999) considered long-lag Phillips curve specifications. In their pseudo out-of-sample results (six inflation measures, four- and eight-quarters ahead, forecast period of 1975–1998), standard Phillips curve forecasts are outperformed by longer-lag versions (25-quarter polynomial distributed lag specifications). Using in-sample statistics, they reject coefficient stability; they attribute the instability to a shift in the NAIRU in the 1990s, not to a change in the slope coefficients in the long-lag specification.

A final paper documenting poor Phillips curve forecasting performance, contemporaneous with Atkeson and Ohanian (2001), is Camba-Mendez and Rodriguez-Palenzuela (2003; originally published as a 2001 European Central Bank working paper). They showed that inflation forecasts at the one-year horizon based on realizable (that is, backward–looking) output gap measures, for the forecast period 1980–1999, underperform the AR benchmark.

In short, during the 1990s a number of papers provided results that activity-based inflation forecasts provided a smaller advantage relative to an AR benchmark after the mid-1980s than these forecasts had before. Ambiguities remained, however, because this conclusion seemed to depend on the sample period and specification, and in any event one could find predictors which were exceptions in the sense that they appeared to provide improvements in the later sample, even if their performance was lackluster in the earlier sample.

Atkeson and Ohanian (2001)

Atkeson and Ohanian (2001) (AO) resolved the ambiguities in this literature from the 1990s by adopting a new, simple univariate benchmark: the forecast of inflation over the next four quarters is the value of four-quarter inflation today.[5] Atkeson and Ohanian showed that this four-quarter random walk forecast improved substantially upon the AR benchmark over the 1984–1999 period. Figure 3.4 plots the moving RMSE of four-quarter ahead forecasts of CPI-all inflation for three univariate forecasts: the AR(AIC) forecast (3), the AO forecast (4), and the UC-SV forecast (5)–(8). Because the AO benchmark improved upon the AR forecast over the 1984–1999 period, and because the AR forecast had more or less the same performance as the unemployment-based Phillips curve on average over this period (see figure 3.2), it is not surprising that the AO forecast

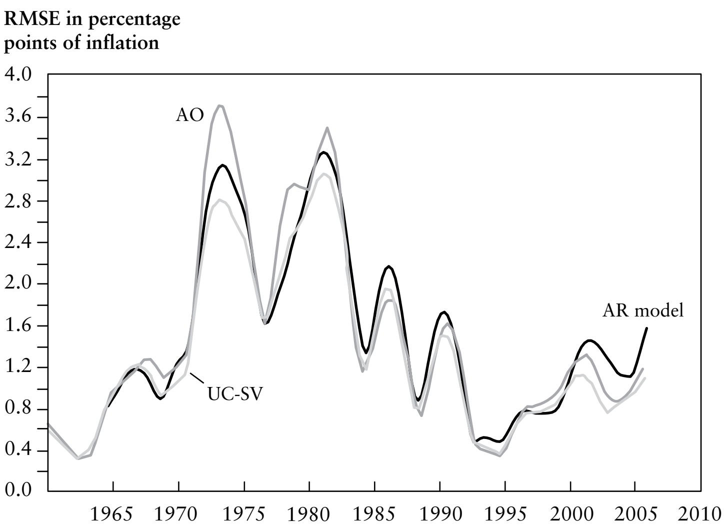


Figure 3.4
Rolling Root Mean Squared Errors for Univariate CPI-All Inflation Forecasts: AR(AIC), Atkeson-Ohanian (AO), and Unobserved Components-Stochastic Volatility (UC-SV) Models

outperformed the Phillips curve forecast over the 1984–1999 period. As Atkeson and Ohanian dramatically showed, across 264 specifications (three inflation measures, CPI-all, CPI-core, and PCE-all, two predictors, the unemployment rate and the Chicago Fed National Activity Index [CFNAI], and various lag specifications), the relative RMSEs of a Phillips curve forecast to the AO benchmark ranged from 0.99 to 1.94: gains from using a Phillips curve forecast were negligible at best, and some Phillips curve forecasts went badly wrong. Atkeson and Ohanian went one step further and demonstrated that, over the 1984–1999 period, Greenbook forecasts of inflation also underperformed their four-quarter random walk forecast.

As figures 3.2 and 3.4 demonstrate, one important source of the problem with Phillips curve forecasts was their poor performance in the second half of the 1990s, a period of strong, but at the time unmeasured, productivity growth that held down inflation. The apparent quiescence of inflation in the face of strong economic growth was puzzling at the time (for example, see Lown and Rich 1997).

An initial response to Atkeson and Ohanian's result was to check whether their claims were accurate; with a few caveats, by and large these were, although only for the post-1984 period they considered. Sims (2002) confirmed Atkeson and Ohanian's results post-1984, but stressed that the AO model performs poorly over the 1979–1983 sample period. Bernanke (2003) cited unpublished work by Board of Governors staff that Atkeson and Ohanian's conclusions do not extend to periods of greater macroeconomic and inflation volatility. Fisher, Liu, and Zhou (2002) used rolling regressions with a 15-year window and showed that Phillips curve models outperformed the AO benchmark in 1977–1984, and also showed that for some inflation measures and some periods the Phillips curve forecasts outperform the AO benchmark post-1984 (for example, Phillips curve forecasts improve upon AO forecasts of PCE-all over 1993–2000). Fisher, Liu, and Zhou (2002) also pointed out that Phillips curve forecasts based on the CFNAI achieve 60–70 percent accuracy in directional forecasting of the change of inflation, compared with 50 percent for the AO coin flip. Fisher, Liu, and Zhou suggested that Phillips curve forecasts do relatively poorly in periods of low inflation volatility and after a regime shift.

Stock and Watson (2003) extended Atkeson and Ohanian's analysis to additional activity predictors (as well as other predictors) and confirmed the dominance of the AO forecast over 1985–1999 at the one-year horizon. Brave and Fisher (2004) extended Atkeson and Ohanian's and Fisher, Liu, and Zhou's (2002) analyses by examining additional predictors and combination forecasts. Brave and Fisher's (2004) findings are broadly consistent with Fisher, Liu, and Zhou (2002) in the sense that they found some individual and combination forecasts that outperform AO over 1993–2000, although not over 1985–1992. Orphanides and van Norden (2005) focused on Phillips curve forecasts using real-time gap measures, and they concluded that although ex post gap Phillips curves fit well using in-sample statistics, when real-time gaps and pseudo out-of-sample methods are used these too improve upon the AR benchmark prior to 1983, but fail to do so over the 1984–2002 sample period.

There are three notable recent studies that confirm Atkeson and Ohanian's basic finding and extend it, with qualifications. First, Stock and Watson (2007) focused on univariate models of inflation and confirmed that the good performance of the AO random walk forecast, relative to other univariate models, is specific to the four-quarter horizon and to Atkeson and Ohanian's sample period. At any point in time, the UC-SV model implies an IMA(1,1) model for inflation, with time-varying coefficients. The forecast function of this IMA(1,1) closely matches the implicit AO forecast function over the 1984–1999 sample—however the models diverge over other subsamples. Moreover, the rolling IMA(1,1) is in turn well approximated by a ARMA(1,1) because the estimated AR coefficient is nearly one.[6] Stock and Watson (2007) also reported some (limited) results for bivariate forecasts using activity indicators (unemployment, one-sided gaps, and output growth) and confirmed Atkeson and Ohanian's finding that these Phillips curve forecasts fail to improve systematically on the AO benchmark or the UC-SV benchmark over the AO sample at the four-quarter horizon.

Second, Canova (2007) undertook a systematic evaluation of four- and eight-quarter ahead inflation forecasts for G7 countries using recursive forecasts over 1996–2000, using a variety of activity variables (unemployment, employment, output gaps, GDP growth) and other indicators (yield curve slope, money growth) as predictors. He found that, for the United States, bivariate direct regressions and trivariate VARs and BVARs did not improve upon the univariate AO forecast, and that there was evidence of instability of forecasts based on individual predictors. Canova (2007) also considered combination forecasts and forecasts generated using a New Keynesian Phillips curve. Over the 1996–2000 U.S. sample, combination forecasts provided a small improvement over the AO forecast, and the New Keynesian Phillips curve forecasts were never the best and generally fared poorly. In the case of the United States, at least, these findings are not surprising in light of the poor performance of Phillips curve forecasts during the low-inflation boom of the second half of the 1990s.

Third, Ang, Bekaert, and Wei (2007) conducted a thorough assessment of forecasts of CPI, CPI-core, CPI excluding housing, and PCE inflation, using 10 variants of Phillips curve forecasts, 15 variants of term structure forecasts, combination forecasts, and ARMA(1,1) and AR(1)-regime switching univariate models in addition to AR and AO benchmarks. They too confirmed Atkeson and Ohanian's basic message that Phillips curve models fail to improve upon univariate models over forecast periods 1985–2002 and 1995–2002. Ang, Bekaert, and Wei's (2007) results constitute a careful summary of the current state of knowledge of inflation forecasting models (both Phillips curve and term structure) in the United States. One finding in their study is that combination forecasts do not systematically improve on individual indicator forecasts, a result that is puzzling in light of the success reported elsewhere of combination forecasts (we return to this puzzle below).[7]

Following Romer and Romer (2000),[8] Sims (2002) and Ang, Bekaert, and Wei (2007) considered professional and survey forecasts, variously including the Federal Reserve Board's Greenbook, Data Resources, Inc., the Michigan Survey of Consumer Sentiment, the Philadelphia Fed's Livingston Survey, the Survey of Professional Forecasters, and Blue Chip surveys. Sims concluded that the Greenbook forecast outperformed the Atkeson-Ohanian forecast over the 1979–1995 period, but not over 1984–1995. Ang, Bekaert, and Wei (2007) found that, for the inflation measures that the survey respondents are asked to forecast, the survey forecasts nearly always beat the ARMA(1,1) benchmark, their best-performing univariate model over the 1985–2002 period; this finding is surprising in light of the literature that has postdated Atkeson and Ohanian (2001). Further study of rolling regressions led Ang, Bekaert, and Wei (2007) to suggest that the relatively good performance of the survey forecasts might be due to the ability of professional forecasters to recognize structural change more quickly than automated regression-based forecasts.[9]

An alternative forecast, so far unmentioned, is that inflation is constant. This forecast works terribly over the full sample but Diron and Mojon (2008) found out that, for PCE-core from 1995:Q1–2007:Q4, a forecast of a constant 2.0 percent inflation rate outperforms AO and AR forecasts at the eight-quarter ahead horizon, although the AO forecast is best at the four-quarter horizon. Diron and Mojon choose 2.0 percent as representative of an implicit inflation target over this period; however, because the United States does not have an explicit ex ante inflation target, this value was chosen retrospectively and this choice does not constitute a pseudo out-of-sample forecast.

The evidence of forecast instability in the foregoing papers is based on changes in relative RMSEs, in some cases augmented by Diebold-Mariano (1995) or West (1996) tests using asymptotic critical values. As a logical matter, the apparent statistical significance of the changes in the relative RMSEs between sample periods could be a spurious consequence of using a poor approximation to the sampling distribution of the relevant statistics. Accordingly, Clark and McCracken (2006) undertook a bootstrap evaluation of the relative RMSEs produced using real-time output gap Phillips curves for forecasting the GDP price deflator and CPI-core. They reached the more cautious conclusion that much of the relatively poor performance of forecasts using real-time gaps could simply be a statistical artifact that is consistent with a stable Phillips curve, although they did find evidence of instability in coefficients on the output gap. One interpretation of the Clark-McCracken (2006) finding is that, over the 1990–2003 period, there are only 14 nonoverlapping observations on the four-quarter ahead forecast error, and estimates of ratios of variances with 14 observations inevitably have a great deal of sampling variability. Rossi and Sekhposyan (2007) also took a careful look at the statistical evidence for breaks using pseudo out-of-sample forecast statistics; theirs is one of the few studies also to use real-time data. Their formal tests for a one-time reversal of forecast performance find a sharp decline in the predictive ability of Phillips curve forecasts post-1984. Additional work is needed to reconcile the results in Clark and McCracken (2006) and Rossi and Sekhposyan (2007).

Attempts to Resuscitate Multivariate Inflation Forecasts, 1999–2007

One response to Atkeson and Ohanian's findings has been to redouble efforts to find reliable multivariate forecasting models for inflation. Some of these efforts used statistical tools, including dynamic factor models, other methods for using a large number of predictors, time-varying parameter multivariate models, and nonlinear time series models. Other efforts exploited restrictions arising from economics, in particular from no-arbitrage models of the term structure. Unfortunately, these efforts have failed to produce substantial and sustained improvements over the AO or UC-SV univariate benchmarks.

Many-predictor forecasts I: dynamic factor models. The plethora of activity indicators used in Philips curve forecasts indicates that there is no single, most natural measure; in fact, these indicators can all be thought of as different measures of overall economic activity. This suggests modeling the activity variables jointly using a dynamic factor model (Geweke 1977, Sargent-Sims 1977), estimating the common latent factor (underlying economic activity), and using that estimated factor as the activity variable in Phillips curve forecasts. Accordingly, Stock and Watson (1999) examined different activity measures as predictors of inflation, estimated (using principal components, as justified by Stock and Watson 2002) as the common factor among 85 monthly indicators of economic activity, and also as the first principal component of 165 series, including the activity indicators plus other series. In addition to using information in a very large number of series, Stock and Watson (2002) showed that principal components estimation of factors can be robust to certain types of instability in a dynamic factor model. Stock and Watson's (1999) empirical results indicated that these estimated factors registered improvements over the AR benchmark and over single-indicator Phillips curve specifications in both 1970–1983 and 1984–1996 subsamples.

A version of the Stock-Watson (1999) common factor, computed as the principal component of 85 monthly indicators of economic activity, has been published in real time since January 2001 as the Chicago Fed National Activity Index (CFNAI). Hansen (2005s) confirmed the main findings in Stock and Watson (1999) about the predictive content of these estimated factors for inflation, relative to a random walk forecast over a forecast period of 1960–2000.

Recent studies, however, have raised questions about the marginal value of Phillips curve forecasts based on estimated factors, such as the CFNAI, for the post-1985 data. As discussed above, Atkeson and Ohanian showed that the AO forecast outperformed CFNAI-based Phillips curves over the 1984–1999 period; this is consistent with Stock and Watson (1999) finding a small improvement in dynamic factor model (DFM) forecasts over this period because Stock and Watson (1999) used an AR benchmark. Banerjee and Marcellino (2006) also found that Phillips curve forecasts using estimated factors perform relatively poorly for CPI-all inflation over a 1991–2001 forecast period. On the other hand, for the longer sample of 1983–2007, Gavin and Kliesen (2008) found that recursive factor forecasts improve upon both the direct AR(12) (monthly data) and AO benchmarks (relative RMSEs are between .88 and .95). In a finding that is inconsistent with Atkeson and Ohanian (2001) and with figure 3.4, Gavin and Kliesen (2008) also found that the AR(12) model outperforms AO at the 12-month horizon for three of the four inflation series; presumably this surprising result is either a consequence of using a slightly different sample than Atkson and Ohanian (in particular, including 1983) or indicates some subtle differences between using quarterly data (as in Atkeson and Ohanian and in figure 3.4) and monthly data.

Additional papers which use estimated factors to forecast inflation include Watson (2003), Bernanke, Bovin, and Elias (2005), Boivin and Ng (2005, 2006), D'Agostino and Giannone (2006), and Giaccomini and White (2006). In an interesting meta-analysis, Eichmeier and Ziegler (2008) considered a total of 52 studies of inflation and/or output forecasts using estimated factors, including 22,849 relative RMSEs for inflation forecasts in the United States and other countries. The dependent variable in their meta-regressions is the RMSE of a factor forecast relative to a benchmark. Eichmeier and Ziegler (2008) concluded that factor model inflation forecasts tend to outperform small model forecasts by a small margin. They also concluded that factor inflation forecasts tend to improve as the horizon increases, and that they improve as the number of series used to estimate the factors increases. Eichmeier and Ziegler's (2008) meta-regressions do not control for sample period, a strategy that permits estimating the average performance of different methods but prevents examining the time-varying relative performance found in the other papers reviewed here. Although Eichmeier and Ziegler (2008) do include indicator variables for the category of benchmark in their meta-regressions, the relative performance of those benchmarks changes over time and this too complicates the interpretation of their results for the purposes of this survey.

Many-predictor forecasts II: Forecast combination, Bayesian Model Averaging, Bagging, and other methods. Other statistical methods for using a large number of predictors are available and have been tried for forecasting inflation. One approach is to use leading index methods, in essence a model selection methodology. In the earliest high-dimensional inflation forecasting exercise of which we are aware, Webb and Rowe (1995) constructed a leading index of CPI-core inflation formed using 7 of 30 potential inflation predictors, selected recursively by selecting indicators with a maximal correlation with one-year ahead inflation over a 48-month window, thereby allowing for time variation. This produced a leading index with time-varying composition that improved upon an AR benchmark over the 1970–1994 period; however, Webb and Rowe (1995) did not provide sufficient information to assess the success of this index post-1983.

A second approach is to use forecast combination methods, in which forecasts from multiple bivariate models (each using a different predictor, lag length, or specification) are combined. Combination forecasts have a long history of success in economic applications—see the review in Timmermann (2006)—and are less susceptible to structural breaks in individual forecasting regressions because, in effect, these combination forecasts average out intercept shifts (Hendry and Clements 2004). Papers that include combination forecasts (pooled over models) include Stock and Watson (1999, 2003), Clark and McCracken (2006), Canova (2007), Ang, Bekaert, and Wei (2007), and Inoue and Kilian (2008). Although combination forecasts often improve upon the individual forecasts, on average these do not substantially improve upon, and are often slightly worse than, factor-based forecasts.

A third approach is to apply model combination or model averaging tools, such as Bayesian Model Averaging (BMA), bagging, and LASSO, developed in the statistics literature for prediction using large data sets. Wright (2003) applied BMA to forecasts of CPI-all, CPI-core, PCE, and the GDP deflator, obtained from 30 predictors, and finds that BMA tended to improve upon simple averaging. Wright's (2003) relative

RMSEs are considerably less than one during the 1987–2003 sample; however this appears to be a consequence of a poor denominator model (an AR(1) benchmark) rather than good numerator models. Inoue and Kilian (2008) considered CPI-all forecasts with 30 predictors using bagging, LASSO, and factor-based forecasts (first principal component), along with BMA, pretest, shrinkage, and some other methods from the statistical literature. They reported a relative RMSE for the single-factor forecast of .80, relative to an AR(AIC) benchmark at the 12 month horizon over their 1983–2003 monthly sample. This is a surprisingly low value in light of Atkeson and Ohanian and subsequent literature, but (like Wright 2003) this low relative RMSE appears to be driven by the use of the AR (instead of AO or UC-SV) benchmark and by the sample period, which includes 1983. Inoue and Kilian (2008) found negligible gains from using the large dataset methods from the statistics literature: the single-factor forecasts beat almost all the other methods they examine, although in most cases the gains from the factor forecasts are slight (the relative RMSEs, relative to the single-factor model, range from .97, for LASSO, to 1.14).

A fourth approach is to model all series simultaneously using high-dimensional VARs with strong parameter restrictions. Bańbura, Gianonne, and Reichlin (2008) performed a pseudo out-of-sample experiment forecasting CPI-all inflation using Bayesian VARs with 3 to more than 100 variables. Over the 1970–2003 sample, they found substantial improvements of medium- to large-dimensional VARs relative to very low-dimensional VARs, but their results are hard to relate to the others in this literature because they do not report univariate benchmarks and do not examine split samples.

In summary, in some cases (some inflation series, some time periods, and some horizons) it appears to be possible to make gains using many predictor methods, either factor estimates or other methods. However, those gains are modest and not systematic and do not substantially overturn Atkeson and Ohanian's (2001) negative results.

Nonlinear models. If the conditional expectation of future inflation is a nonlinear function of the predictors, and if the predictors are persistent, then linear approximations to the conditional mean function can exhibit persistent time variation. Thus the time variation documented above could be a consequence of using linear models. Accordingly, one approach to the apparent time variation in the inflation-output relation is to consider nonlinear Phillips curves and nonlinear univariate time series models. There is a substantial literature on nonlinear Phillips curves that reports only in-sample measures of fit, not pseudo out-of-sample forecasts; see Dupasquier and Ricketts (1998) and Barnes and Olivei (2003) for references. Barnes and Olivei (2003) is a noteworthy paper in that literature; they consider a piecewise linear specification and use dynamic simulations to argue that this specification (with a time-varying NAIRU) provides a better description of the late 1990s and early 2000s than does a linear specification. Their specification is capable of producing the episodically effective Phillips curve forecasts seen in the forecasting literature. Yet their use of only in-sample statistics makes it difficult to compare their findings to the forecasting literature that is the focus of this survey. Papers that evaluate nonlinear inflation forecasting models using pseudo out-of-sample methods include Dupasquier and Ricketts (1998), Moshiri and Cameron (2000), Tkacz (2000), Ascari and Marrocu (2003), and Marcellino (2008).

We read the conclusions of this literature on nonlinear Phillips curves and nonlinear univariate time series models as negative. Although nonlinearities are found using in-sample statistics, the pseudo out-of-sample literature fails to confirm any benefits of nonlinear models for forecasting inflation. Marcellino (2008) examined univariate rolling and recursive CPI-all forecasts (over 1980–2004 and 1984–2004) using logistic smooth transition autoregressions and neural networks (a total of 28 nonlinear models) and found little or no improvement from using nonlinear models. He also documented that nonlinear models can produce outlier forecasts, presumably because of overfitting. Ascari and Marrocu (2003) and Moshiri and Cameron (2000), who apply artificial neural networks to Canadian data, also provided negative conclusions. These negative results in the pseudo out-of-sample literature mean that exploitable nonlinearities have not been found, but not that they do not exist. Indeed, the in-sample results of Barnes and Olivei (2005) presage findings reported below in section 5.

Structural term structure models. Until now, this survey has concentrated on forecasts from the first two families of inflation forecasts (prices-only and Phillips curve forecasts). One way to construct inflation forecasts in the third family—forecasts based on forecasts made by others—is to make inflation forecasts using the term structure of interest rates, as in (11). Starting with Barsky (1987), Mishkin (1990a, 1990b, 1991), and Jorion and Mishkin (1991), there is a large literature that studies such forecasting regressions. The findings of this literature, which are reviewed in Stock and Watson (2003), are generally negative; that is, term spread forecasts do not improve over Phillips curve forecasts in the pre-1983 period, and they do not improve over a good univariate benchmark in the post-1984 period.

This poor performance of first-generation term spread forecasts is evident in figure 3.5, which plots the rolling RMSE of the pseudo out-of-sample forecast based on the recursively estimated term spread model (11), along with the RMSEs of the AR(AIC) and AO univariate benchmarks. Term spreads are typically one of the variables included in the forecast comparison studies discussed earlier (Fisher, Liu, and Zhou 2002, Canova 2007, and Ang, Bekaert, and Wei 2007) and these recent studies also reach the same negative conclusion about unrestricted term spread forecasting regressions, either as the sole predictor or when used in addition to an activity indicator.

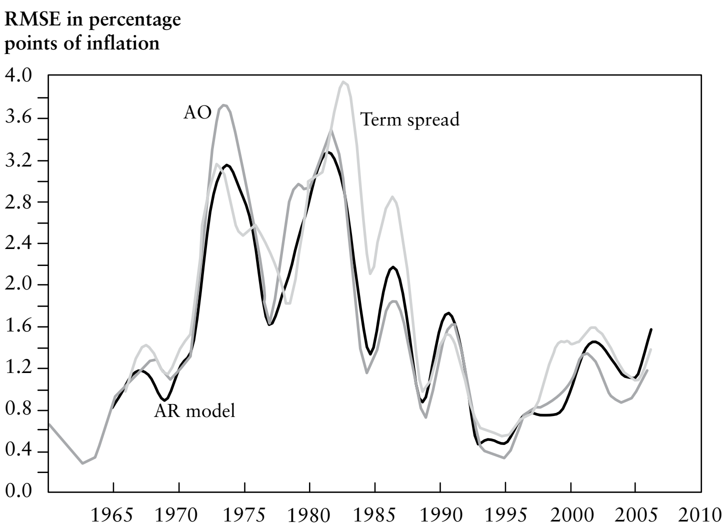


Figure 3.5
Rolling Root Mean Squared Errors for CPI-All Inflation Forecasts: AR(AIC), Atkeson-Ohanian (AO), and Term Spread Model (ADL-*spread*)

Recent attempts to forecast inflation using term spreads have focused on employing economic theory, in the form of no-arbitrage models of the term structure, to improve upon the reduced-form regressions, such as (11). Most of this literature uses full-sample estimation and measures of fit; see Ang, Bekaert, and Wei (2007), DeWachter and Lyrio (2006), and Berardi (2007) for references. The one paper of which we are aware that produces pseudo out-of-sample forecasts of inflation is Ang, Bekaert, and Wei (2007), who considered four-quarter ahead forecasts of CPI-all, CPI-core, CPI-excluding housing, and PCE inflation using two no-arbitrage term structure models, one with constant coefficients and one with regime switches. Neither model forecasted well, with relative RMSEs (relative to an ARMA(1,1)) ranging from 1.05 to 1.59 for the four inflation series and two forecast periods (1985–2002 and 1995–2002).

We are not aware of any papers that evaluate the performance of inflation forecasts backed out of the TIPS yield curve, and such a study would be of considerable interest.

Forecasting using the cross-section of prices. Another approach is to try to exploit information in the cross-section of inflation indexes (percentage growth of sectoral or commodity group price indexes) for forecasting headline inflation. Hendry and Hubrich (2007) used four high-level subaggregates to forecast CPI-all inflation. They explored several approaches, including combining disaggregated univariate forecasts and using factor models. Hendry and Hubrich (2007) found that exploiting the disaggregated information directly to forecast the aggregate improves modestly over an AR benchmark in their pseudo out-of-sample forecasts of CPI-all over 1970–1983 but negligibly over the AO benchmark over 1984–2004 at the 12-month horizon; however they also found that no single method for using the subaggregates works best. If one uses heavily disaggregated inflation measures, then some method must be used to control parameter proliferation, such as the methods used in the many-predictor applications discussed above. In this vein, Hubrich (2005) presented negative results concerning the aggregation of components forecasts for forecasting the Harmonized Index of Consumer Prices in the euro-zone. Reis and Watson (2007) estimated a dynamic factor using a large cross-section of inflation rates but did not conduct any pseudo out-of-sample forecasting.

Rethinking the notion of core inflation suggests different approaches to using the inflation subaggregates. Building on the work of Bryan and Cecchetti (1994), Bryan, Cecchetti, and Wiggins (1997) suggested constructing core inflation as a trimmed mean of the cross-section of prices, where the trimming was chosen to provide the best (in-sample) estimate of underlying trend inflation (measured variously as a 24- to 60-month centered moving average). Smith (2004) investigated the pseudo out-of-sample forecasting properties of trimmed mean and median measures of core inflation (forecast period 1990–2000). Smith (2004) reported that the inflation forecasts based on weighted-median core measures have relative RMSEs of .85 for CPI-all and .80 for PCE-all, relative to an exponentially-declining AR benchmark (she does not consider the AO benchmark), although oddly she found that the trimmed mean performed worse than the AR benchmark.

4. A Quantitative Recapitulation: Changes in Univariate and Phillips Curve Inflation Forecast Performance

This section undertakes a quantitative summary of the literature review in the previous section by considering the pseudo out-of-sample performance of a range of inflation forecasting models using a single consistent data set. The focus is on activity-based inflation forecasting models, although some other predictors are considered. We do not consider survey forecasts or inflation expectations implicit in the TIPS yield curve. As Romer and Romer (2000), Sims (2002), and Ang, Bekaert, and Wei (2007) showed, Greenbook and some median survey forecasts perform quite well and thus are useful for policy work—but our task is to understand how to improve upon forecasting systems, not to delegate this work to others.

Forecasting Models

Univariate models. The univariate models consist of the AR(AIC), Atkeson-Ohanian, and UC-SV models in section 2.2; direct AR models with a fixed lag length of four lags, (AR(4)) and Bayes Information Criterion lag selection (AR(BIC)); and iterated AR(AIC), MA(1), and AR(24) models, where the AR(24) model imposes the Gordon (1990) step function lag restriction and the unit root in π_t. AIC and BIC model selection used a minimum of 0 and a maximum of six lags. Both rolling and recursively estimated versions of these models are considered. In addition some fixed-parameter models were considered: MA(1) models with fixed MA coefficients of 0.25 and 0.65 (these are taken from Stock and Watson 2007), and the monthly MA model estimated by Nelson and Schwert (1977), temporally aggregated to quarterly data (see Stock and Watson 2007, equation (7)).

Triangle and Time-Varying NAIRU models. Four triangle models are considered: specification (9), the results of which were examined in section 3; specification (9) without the supply shock variables (relative price of food and energy, import prices, and Nixon dummies); and these two versions with a time-varying (TV) NAIRU. The TV-NAIRU specification introduces random walk intercept drift into (9) following Staiger, Stock, and Watson (1997) and Gordon (1998); specifically, the TV-NAIRU version of (9) is

$$(12)\quad \pi_{t+1} = \alpha^G(\mathrm{L})\pi_t + \beta(\mathrm{L})(u_{t+1} - \bar{u}_t) + \gamma(\mathrm{L})z_t + v_{t+1},$$

$$(13)\quad \bar{u}_{t+1} = \bar{u}_t + \eta_{t+1},$$

where v_t and η_t are modeled as independent i.i.d. normal errors with relative variance σ^2_η/σ^2_v (recall that $\alpha^G(1) = 1$ so a unit root is imposed in (12)). For the calculations here, σ^2_η/σ^2_v is set to 0.1.

ADL Phillips curve models. The ADL Phillips curve models are direct models of the form,

$$(14)\ \pi^h_{t+h} - \pi_t = \mu^h + \alpha^h(\mathrm{L})\Delta\pi_t + \beta^h(\mathrm{L})x_t + v^h_{t+h},$$

where x_t is an activity variable (an output gap, growth rate, or level, depending on the series). Lag lengths for π_t and x_t are chosen separately by AIC and, alternatively, BIC.

ADL models using other predictors. ADL models are specified and estimated the same way as the ADL Phillips curve model (14), but the activity variable x_t is replaced by another predictor (term spreads, core inflation, and so on).

Combination forecasts. Let $\{\hat{\pi}^h_{i,t+h|t}\}$ denote a set of n forecasts of π^h_{t+h}, made using data through date t. Combined forecasts are computed in three ways: by "averaging" (mean, median, and trimmed mean); by a MSE-based weighting scheme; or by using the forecast that is most recently best. The MSE-based combined forecasts f_t are of the form $f_t = \sum_{i=1}^{n} \lambda_{it} \hat{\pi}^h_{i,t+h|t}$, where six methods are used to compute the weights $\{\lambda_{it}\}$:

(15) (A) $\lambda_{it} = (1/\hat{\sigma}^2_{it}) / \sum_{j=1}^{n} (1/\hat{\sigma}^2_{jt})$, with $\hat{\sigma}^2_{it} = \sum_{j=0}^{39} 0.9^j e^2_{i,t-j}$,

(16) (B) $\lambda_{it} = (1/\hat{\sigma}^2_{it}) / \sum_{j=1}^{n} (1/\hat{\sigma}^2_{jt})$, with $\hat{\sigma}^2_{it} = \sum_{j=0}^{39} 0.95^j e^2_{i,t-j}$,

(17) (C) $\lambda_{it} = (1/\hat{\sigma}^2_{it}) / \sum_{j=1}^{n} (1/\hat{\sigma}^2_{jt})$, with $\hat{\sigma}^2_{it} = \sum_{j=0}^{39} e^2_{i,t-j}$,

(18) (D) $\lambda_{it} = (1/\hat{\sigma}^2_{it})^2 / \sum_{j=1}^{n} (1/\hat{\sigma}^2_{jt})^2$, with $\hat{\sigma}^2_{it} = \sum_{j=0}^{39} 0.9^j e^2_{i,t-j}$,

(19) (E) $\lambda_{it} = (1/\hat{\sigma}^2_{it})^2 / \sum_{j=1}^{n} (1/\hat{\sigma}^2_{jt})^2$, with $\hat{\sigma}^2_{it} = \sum_{j=0}^{39} 0.95^j e^2_{i,t-j}$,

(20) (F) $\lambda_{it} = (1/\hat{\sigma}^2_{it})^2 / \sum_{j=1}^{n} (1/\hat{\sigma}^2_{jt})^2$, with $\hat{\sigma}^2_{it} = \sum_{j=0}^{39} e^2_{i,t-j}$,

where $e_{i,t} = \pi^h_t - \hat{\pi}^h_{i,t|t-h}$ is the pseudo out-of-sample forecast error for the i^{th} h-step ahead forecast and the MSEs are estimated using a 10-year rolling window and, for methods (A), (B), (D), and (E), discounting.

Inverse MSE weighting (based on population MSEs) is optimal if the individual forecasts are uncorrelated, and methods (A) – (C) are different ways to implement inverse MSE weighting. Methods (D) – (F) give greater weight to better-performing forecasts than does inverse MSE weighting. Optimal forecast combination using regression weights as in Bates and Granger (1969) is not feasible with the large number of forecasts under consideration. As Timmerman (2006) notes, equal-weighting (mean combining) often performs well and Timmerman (2006) provides a discussion of when mean combining is optimal under squared error loss.

The "recent best" forecasts are the forecasts from the model that has the lowest cumulative MSE over the past four (or, alternatively, eight) quarters.

Finally, in an attempt to exploit the time-varying virtues of the UC-SV and triangle models, the recent best is also computed using only the UC-SV and triangle model (with time varying NAIRU and *z* variables).

The complete description of models considered is given in the notes to table 3.1.

Results

The pseudo out-of-sample forecasting performance of each forecasting procedure (model and combining method) is summarized in tabular and graphical form.

The tabular summary consists of relative RMSEs of four-quarter ahead inflation forecasts, relative to the UC-SV benchmark, for six forecast periods; these are tabulated in tables 3.1–3.5 for the five inflation series. The minimum model estimation sample was 40 quarters, and blank cells in the table indicate that for at least one quarter in the forecast period there were fewer than 40 observations available for estimation.

The graphical summary of each model's performance is given is figures 3.6–3.11 for the five inflation series. Figure 3.6 presents the rolling RMSE for the UC-SV benchmark for the five inflation series, and figures 3.7–3.11 show the RMSE of the various forecasts relative to the UC-SV benchmark. Part (a) of figure 3.7–3.11 displays the rolling relative RMSE

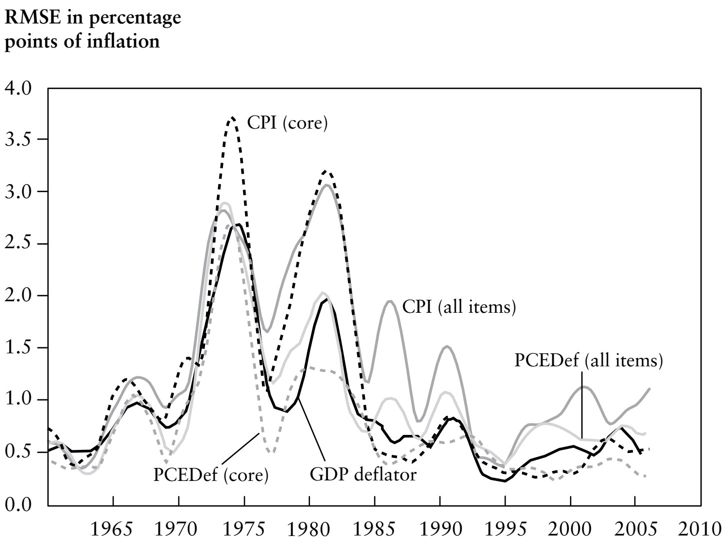


Figure 3.6
Rolling Root Mean Squared Errors for Inflation Forecasts, Unobserved Components-Stochastic Volatility Model, for All Five Inflation Series

for the prototype models, where the rolling RMSE for each model is computed using (2). Parts (b) – (d) plot the ratio of the rolling RMSE for each category of models, relative to the UC-SV model: univariate models in part (b), Phillips curve forecasts (ADL and triangle) in part (c), and combination forecasts in part (d). In each of parts (b) – (d), leading case models or forecasts are highlighted. The unlabeled relative RMSE paths, which are presented using small dots in panels (b)–(d) of figures 3.7–3.11, portray the rolling RMSEs of all other forecasting models in tables 3.1–3.5 for the relevant inflation series and the indicated category of forecast. For example, figure 3.7(c) represents the relative rolling RMSEs for all the Phillips curve forecasts listed in table 3.1, three of which are labeled in the figure while the rest remain unlabeled.

These tables and figures present a great many numbers and facts. Inspection of these results leads us to the following conclusions:

1. There is strong evidence of time variation in the inflation process, in predictive relations, and in Phillips curve forecasts. This is consistent with the literature review, in which different authors reach different conclusions about Phillips curve forecasts depending on the sample period.

2. The performance of Phillips curve forecasts, relative to the UC-SV benchmark, has a considerable systematic component (part (c) of figures 3.7–3.11): during periods in which the ADL-*u* prototype model is forecasting well, reasonably good forecasts can be made using a host of other activity variables. In this sense, the choice of activity variable is secondary to the choice of whether one should use an activity-based forecast.

3. Among the univariate models considered here, with and without time-varying coefficients, there is no single model, or combination of univariate models, that has uniformly better performance than the UC-SV model. Of the 82 cells in table 3.1 that give relative RMSEs for univariate CPI-all forecasts in different subsamples, only four cells have RMSEs less than 1.00, the lowest of which is .95, and these instances are for fixed-parameter MA models in the 1960s and in the 1985–1992 period. Similar results are found for the other four inflation measures. In some cases, the AR models do quite poorly relative to UC-SV. For example, in the 2001–2007 sample the AR forecasts of CPI-all and PCE-all inflation have very large relative MSEs (typically exceeding 1.3). In general, the performance of the AR model, relative to the UC-SV or AO benchmarks, is series- and period-specific. This reinforces the remarks in the literature

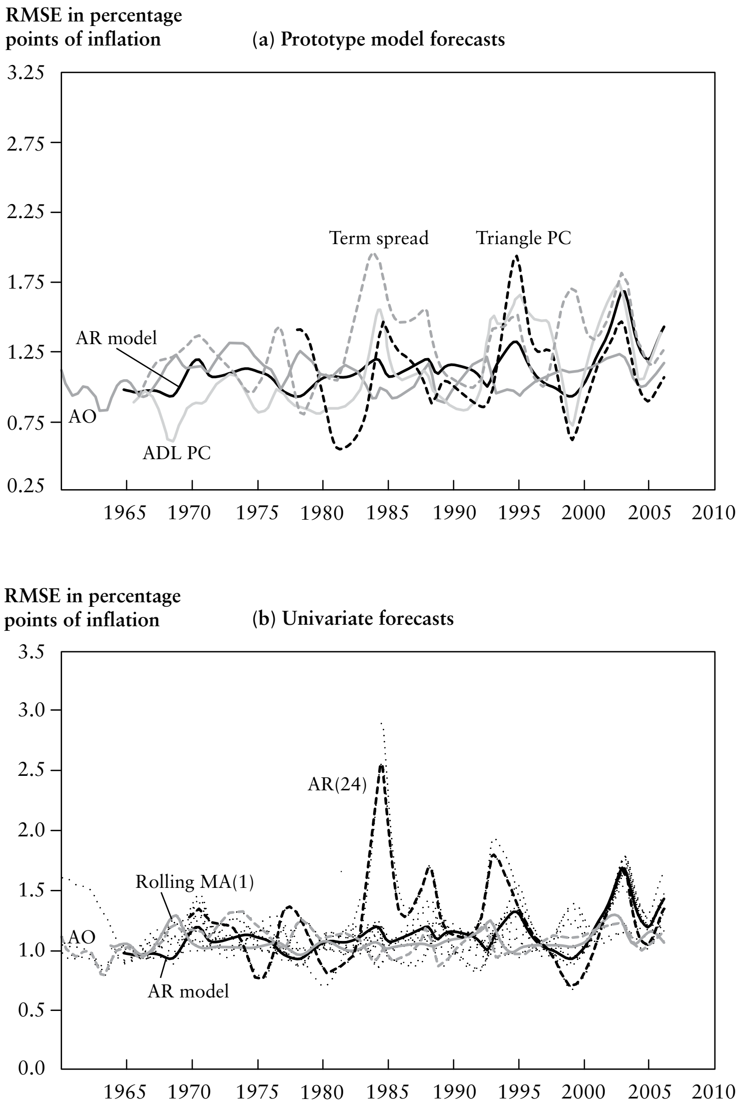
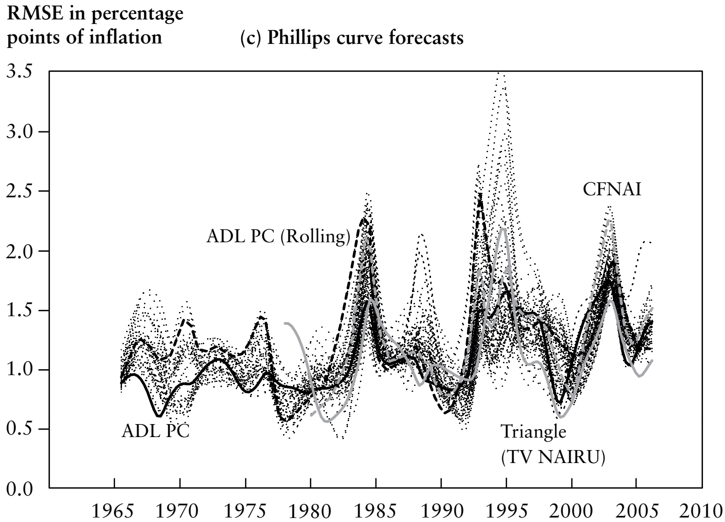


Figure 3.7
Rolling Root Mean Squared Errors, Relative to Unobserved Components-Stochastic Volatility Model: CPI-All



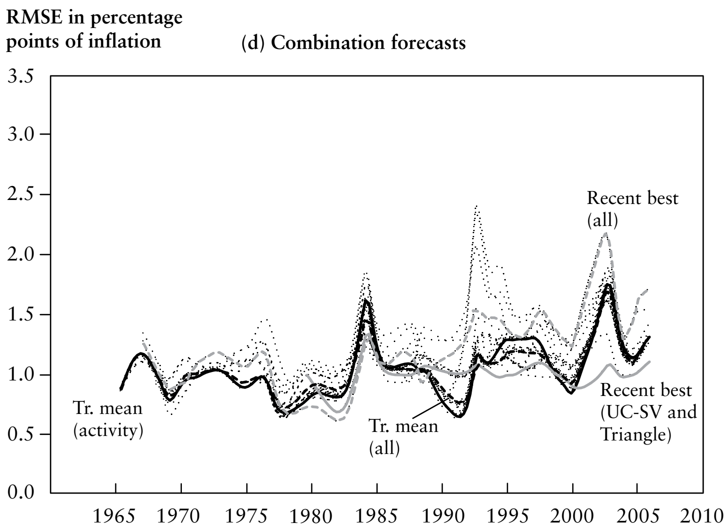
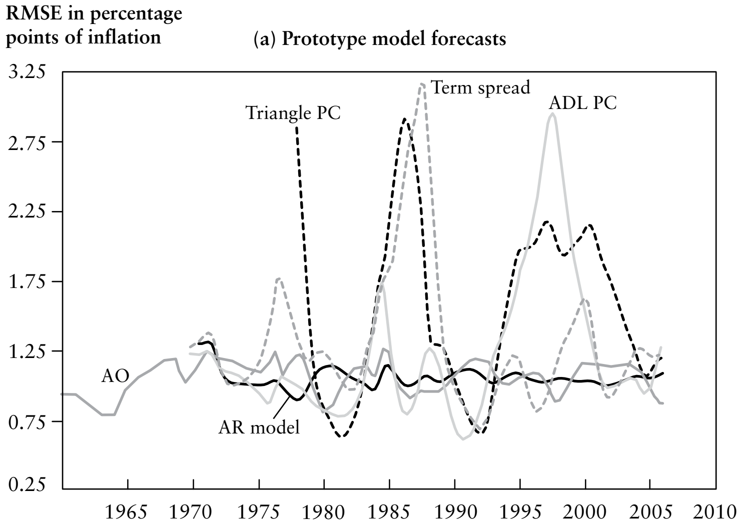


Figure 3.7 (continued)



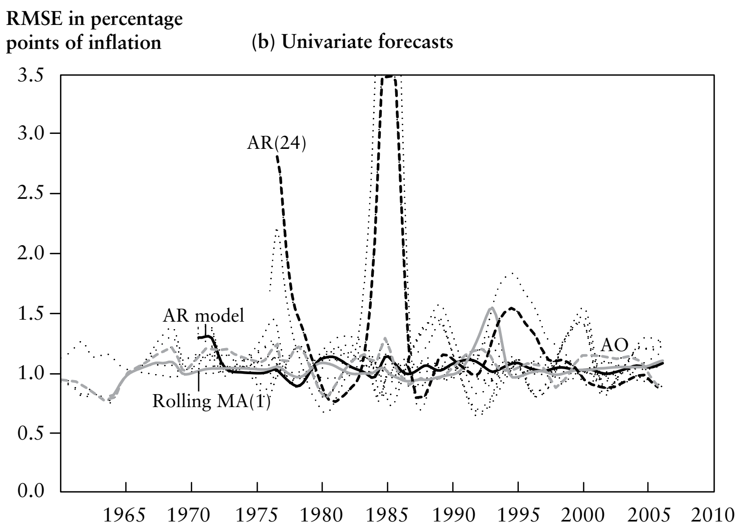
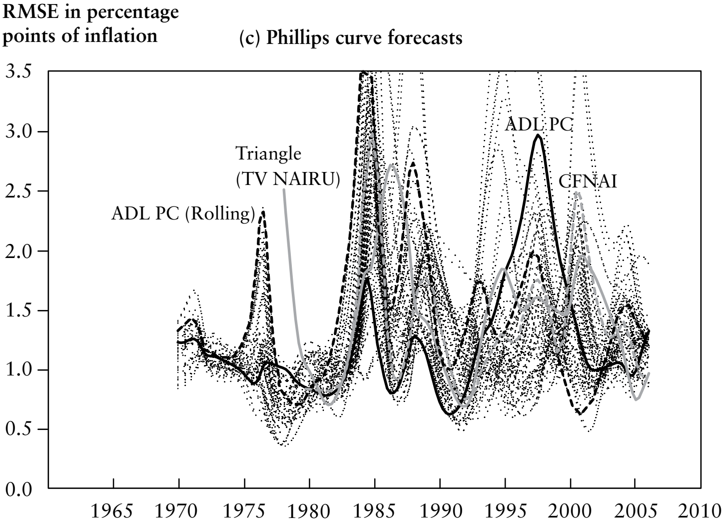


Figure 3.8
Rolling Root Mean Squared Errors, Relative to Unobserved Components-Stochastic Volatility Model: CPI-Core



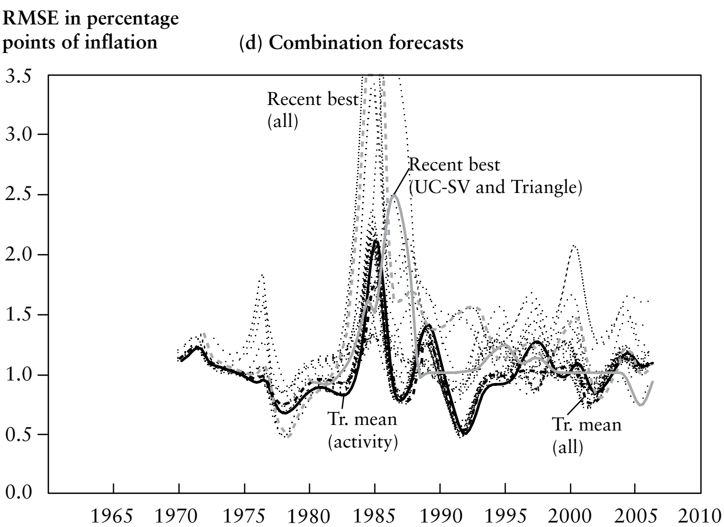
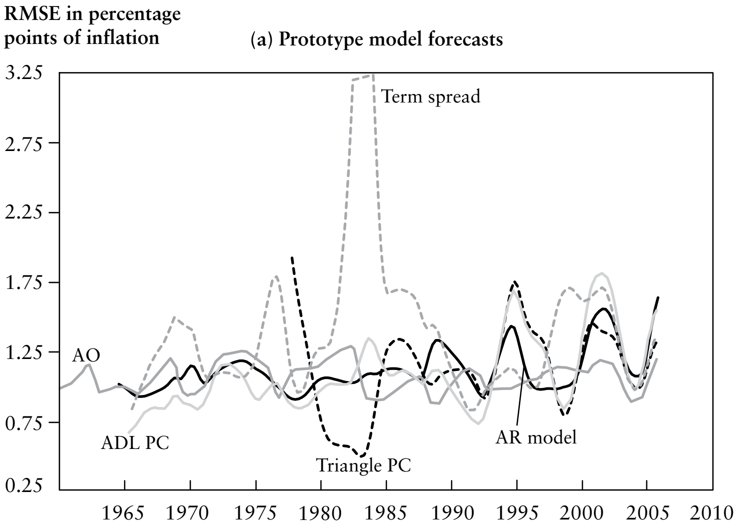


Figure 3.8 (continued)



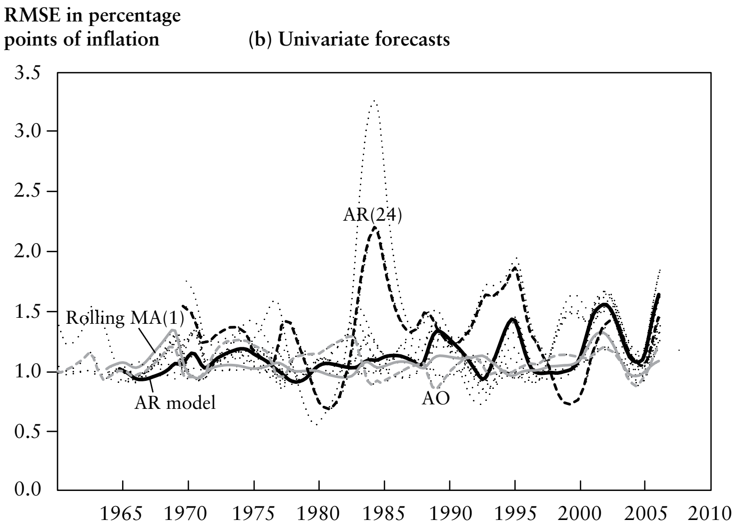
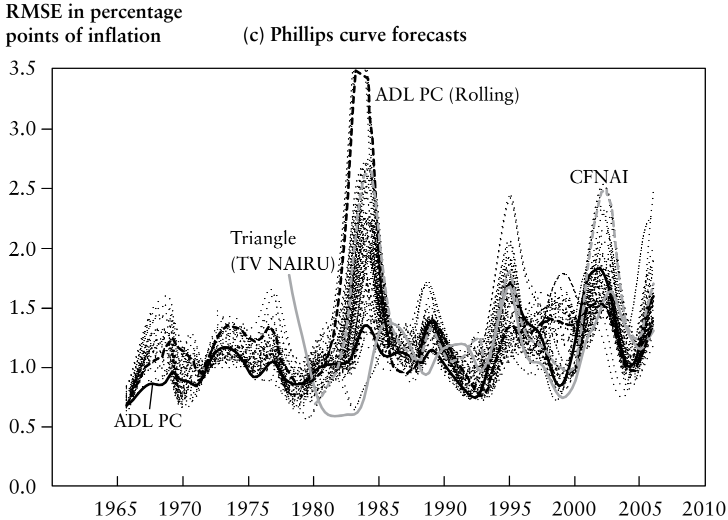


Figure 3.9
Rolling Root Mean Squared Errors, Relative to Unobserved Components-Stochastic Volatility Model: PCE-All



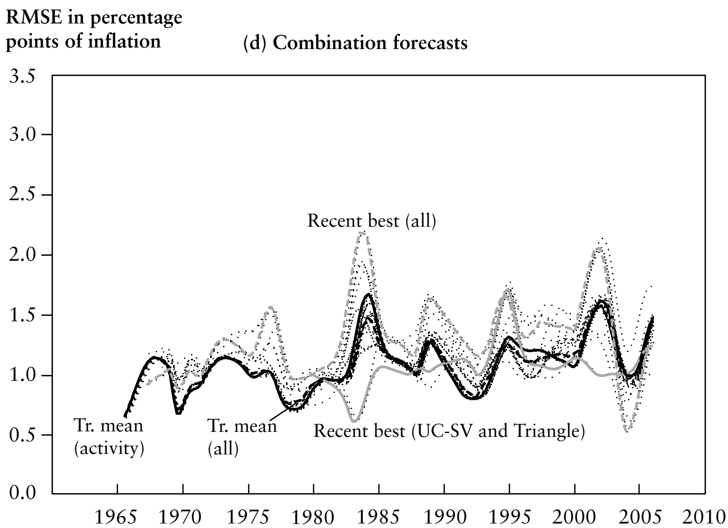


Figure 3.9 (continued)

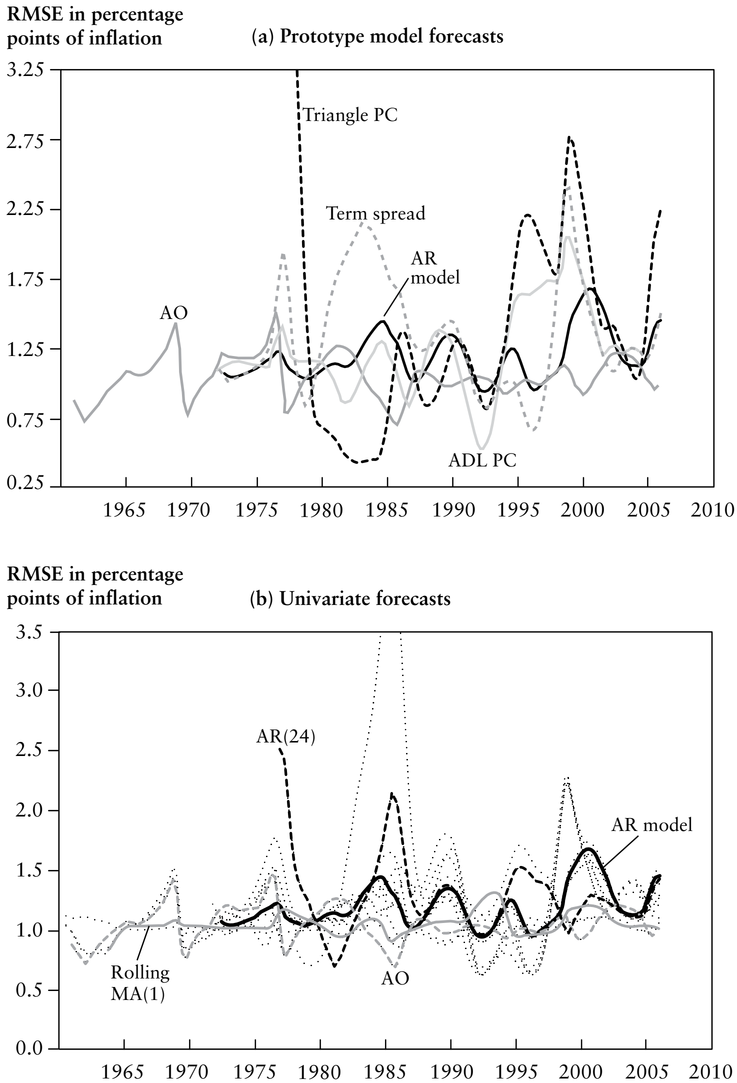


Figure 3.10
Rolling Root Mean Squared Errors, Relative to Unobserved Components-Stochastic Volatility Model: PCE-Core

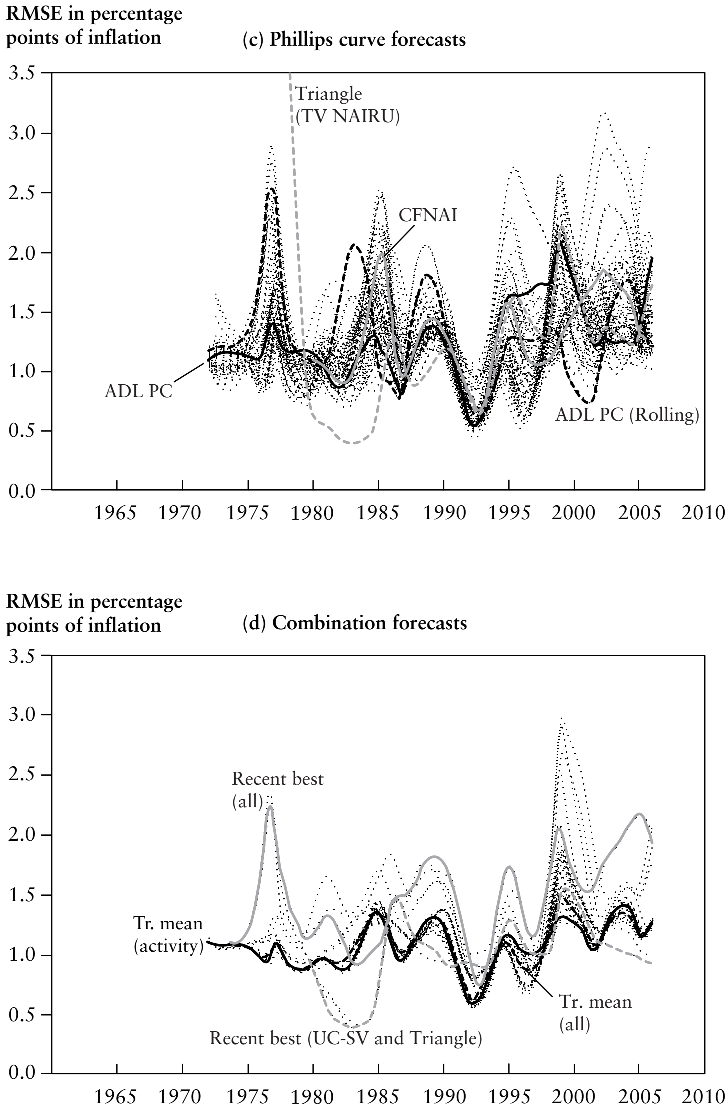


Figure 3.10 (continued)

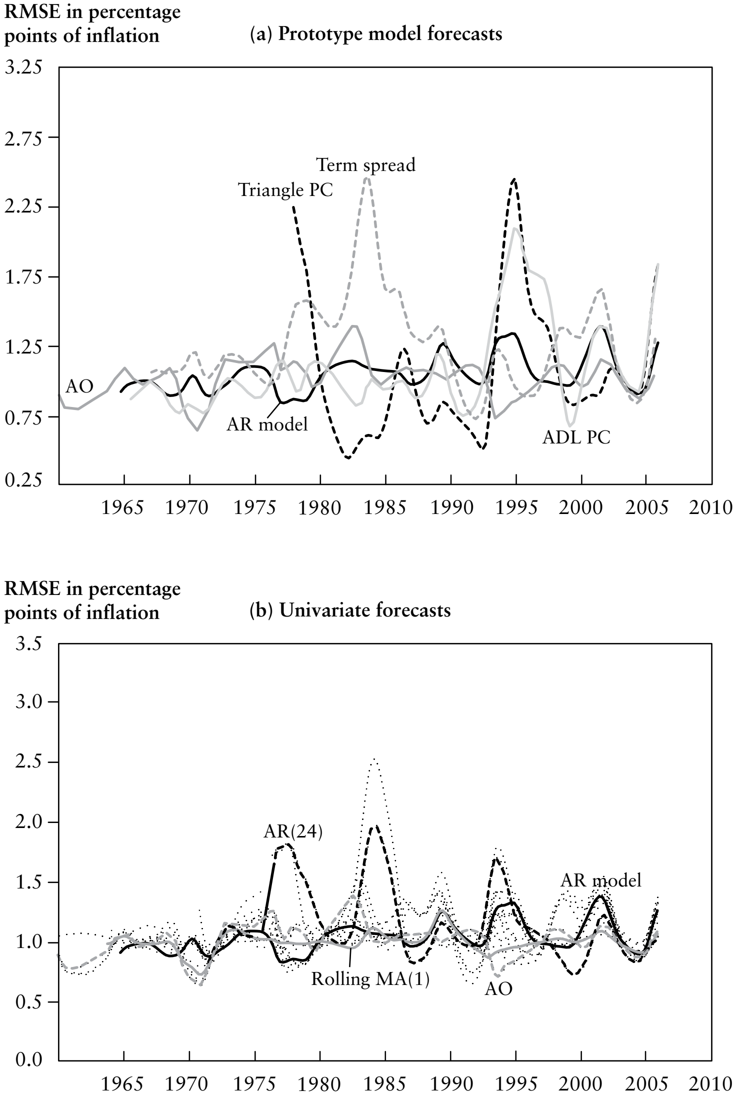


Figure 3.11
Rolling Root Mean Squared Errors, Relative to Unobserved Components-Stochastic Volatility Model: GDP Deflator

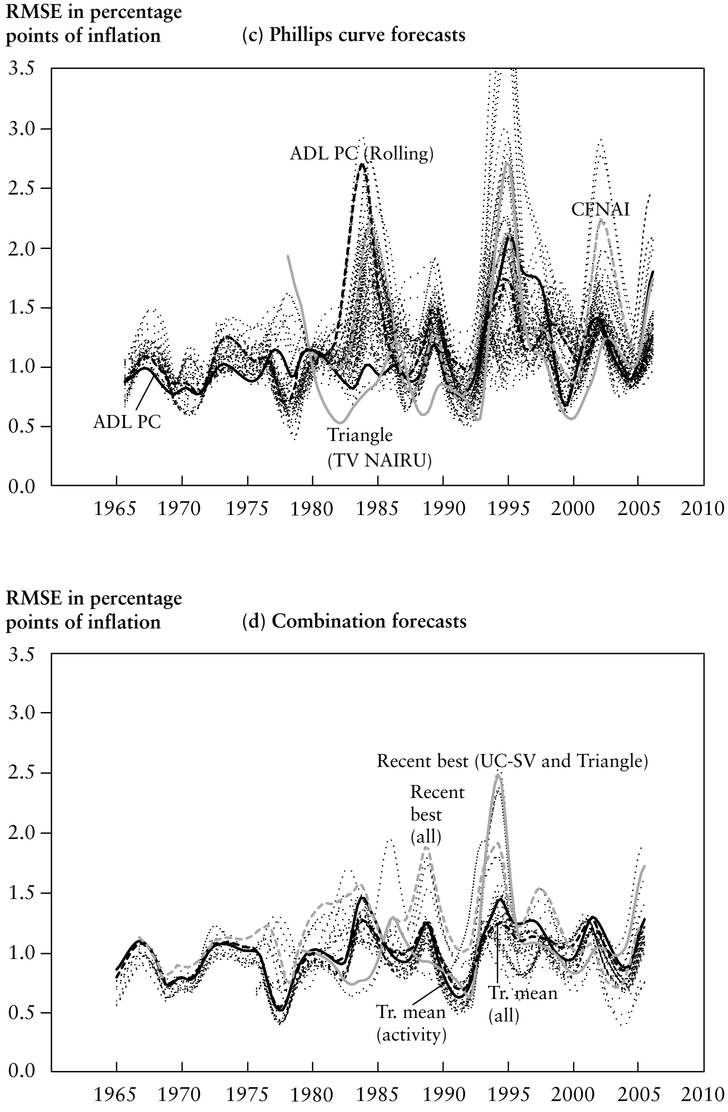


Figure 3.11 (continued)

review about the importance of using a consistently good benchmark: the apparently good performance of a predictor for a particular inflation series over a particular period can be the result of a large denominator, not a small numerator.

4. Although some of the Phillips curve forecasts improved substantially on the UC-SV model during the 1970s and early 1980s, there is little or no evidence that it is possible to improve upon the UC-SV model on average over the full later samples. Nevertheless, there are notable periods and inflation measures for which Phillips curve models do quite well. The triangle model does particularly well during the high unemployment disinflation of the early 1980s for all five inflation measures. For CPI-all, PCE-all, and the GDP deflator, it also does well in the late 1990s, while for CPI-core and PCE-core the triangle model does well emerging from the 1990 recession. This episodically good behavior of the triangle model, and of Phillips-curve forecasts more generally, provides a more nuanced interpretation of the history of inflation forecasting models than the blanket Atkeson-Ohanian (2001) finding which, as stated in their paper's abstract, concluded that "none of the NAIRU forecasts is more accurate than the naïve forecast."

5. Forecast combining, which has worked so well in other applications (Timmerman 2006), generally improves upon the individual Phillips-curve forecasts; however, the combination forecasts generally do not improve upon the UC-SV benchmark in the post-1993 periods. For example, for CPI-all, the mean-combined ADL-activity forecasts have a relative RMSE of .86 over 1977–1982 and .96 over 1985–1992; these mean-combined forecasts compare favorably to individual activity forecasts and to the triangle model. In the later periods, however, the forecasts being combined have relative RMSEs exceeding 1.0; combining them works no magic and fails to improve upon the UC-SV benchmark. Although some of the combining methods improve upon equal weighting, these improvements are neither large nor systematic. In addition, consistent with the results in Fisher, Liu, and Zhou (2002), factor forecasts (using the CFNAI) fail to improve upon the UC-SV benchmark on average over the later periods. These results are consistent with the lack of success found by attempts in the literature (before and after Atkeson and Ohanian 2001) to obtain large gains by using many predictors and/or model combinations.

6. Forecasts using predictors other than activities variables, while not the main focus of this paper, generally fare poorly, especially during the post-

1992 period. For example, the relative RMSE of the mean-combined forecast using nonactivity variables is at least 0.99 in each subsample in tables 3.1–3.5 (23 cases). We did not find substantial improvements using alternative measures of core (median and trimmed mean CPI) as predictors.[10] Although our treatment of nonactivity variables is not comprehensive, these results largely mirror those in the literature.

5. When Were Phillips Curve Forecasts Successful, and Why?

If the relative performance of Phillips curve forecasts has been episodic, is it possible to characterize what makes for a successful or unsuccessful episode?

The relative RMSEs of the triangle and ADL-*u* model forecasts for headline inflation (CPI-all, PCE-all, and GDP deflator), relative to the UC-SV benchmark, are plotted in figure 3.12, along with the unemployment rate. One immediately evident feature is that the triangle model has substantially larger swings in performance than the ADL-*u* model. This said, the dates of relative success of these Phillips curve forecasts bear considerable similarities across models and inflation series. Both models perform relatively well for all series in the early 1980s, in the early 1990s, and around 1999; both models perform relatively poorly around 1985 and in the mid-1990s. These dates of relative success correspond approximately to dates of different phases of U.S. business cycles.

Figure 3.13 is a scatterplot of the quarterly relative RMSE for the triangle (panel a) and ADL-*u* (panel b) prototype models, versus the two-sided unemployment gap (the two-sided gap was computed using the two-sided version of the lowpass filter described in section 2), along with kernel regression estimates. The most striking feature of these scatterplots is that the relative RMSE is minimized, and is considerably less than 1.0, at the extreme values of the unemployment gap, both positive and negative. (The kernel regression estimator exceeds 1.0 at the most negative values of the unemployment gap for the triangle model in panel (a), but there are few observations in that tail.) When the unemployment rate is near the NAIRU (as measured by the lowpass filter), both Phillips curve models do worse than the UC-SV model. But when the unemployment

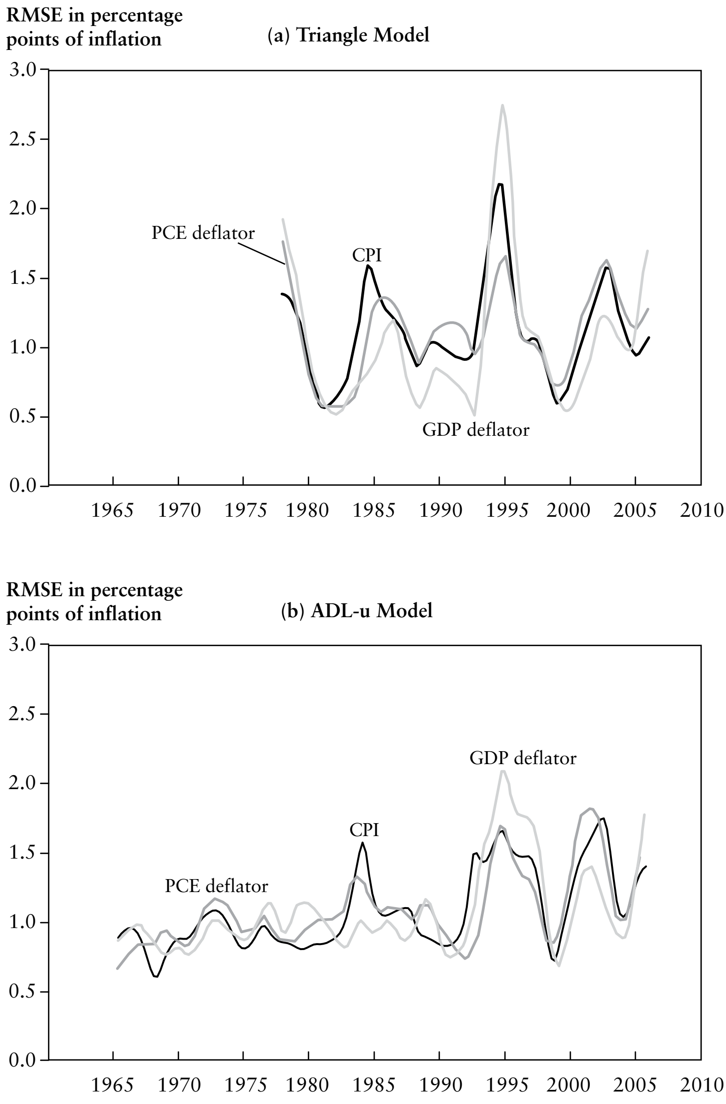


Figure 3.12
Rolling Root Mean Squared Errors, Relative to Unobserved Components-Stochastic Volatility Model, and the Unemployment Rate

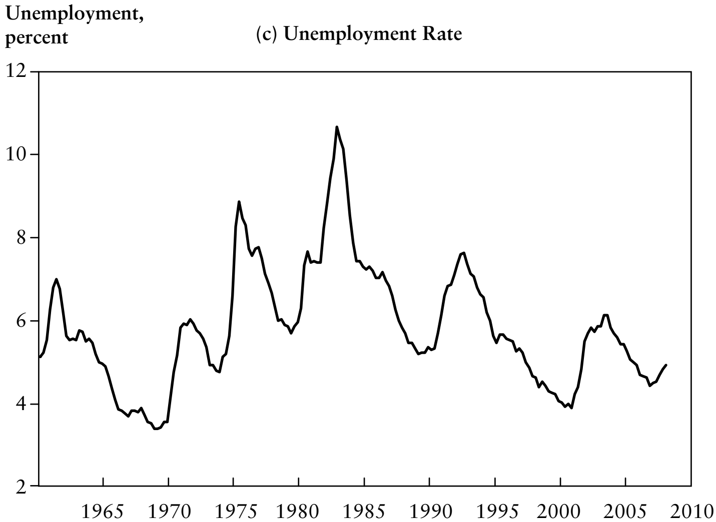


Figure 3.12 (continued)

gap exceeds 1.5 percentage points in absolute value, the Phillips curve forecasts improve substantially upon the UC-SV model. Because the gap is largest in absolute value around turning points, the Phillips curve models provide improvements over the UC-SV model around cyclical turning points, but not during normal times.

Figure 3.14 takes a different perspective on the link between performance of the Phillips curve forecasts and the state of the economy, by plotting the relative RMSE against the four-quarter change in the unemployment rate. The relative improvements in the Phillips curve forecasts do not seem as closely tied to the change in the unemployment rate as to the gap (the apparent improvement at very large changes of the unemployment rate is evident in only a few observations).

Figures 3.15–3.17 examine a conjecture in the literature, that Phillips curve forecasts are relatively more successful when inflation is volatile, by plotting the rolling relative RMSE against the four-quarter change in four-quarter inflation. These figures provide only limited support for this conjecture, as do similar scatterplots (not provided here) of the rolling RMSE against the UC-SV estimate of the instantaneous variance of the first difference of the inflation rate. It is true that the worst performance occurs when in fact inflation is changing very little but, other than for the

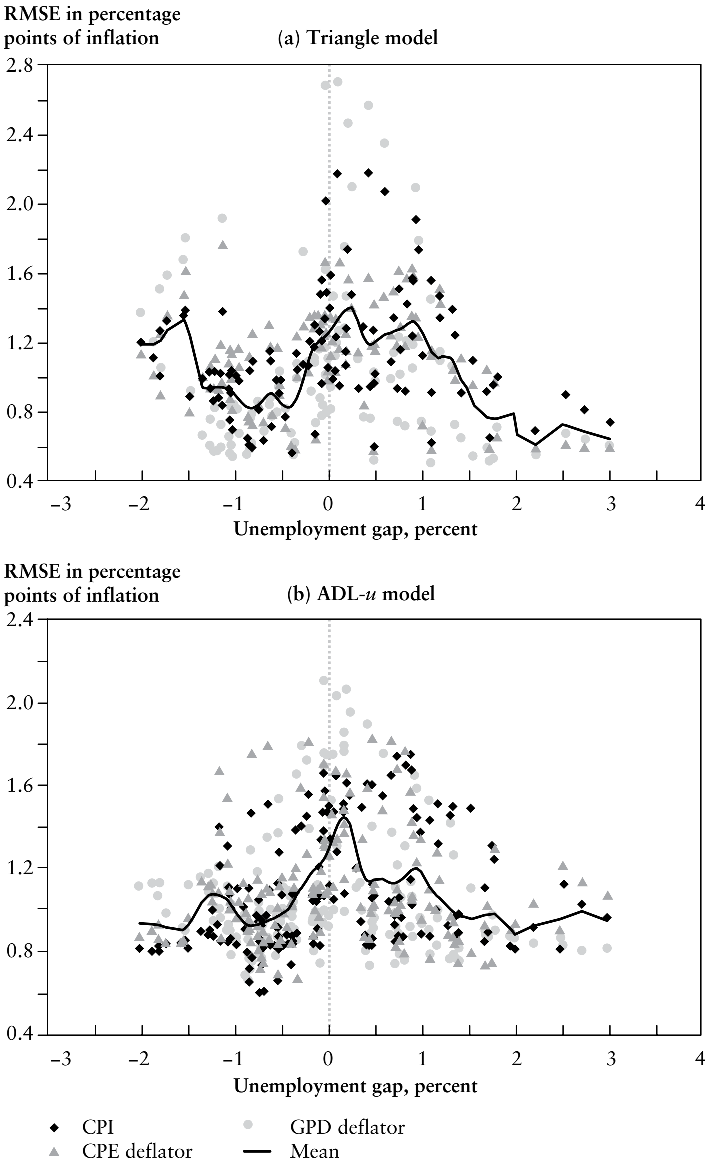


Figure 3.13
Scatterplot of Rolling Root Mean Square Errors of Headline Inflation Forecasts, Relative to Unobserved Components-Stochastic Volatility, vs. the Unemployment Gap (two-sided bandpass)
Note: Mean is kernel regression estimate using data for all three series. Each point represents a quarter.

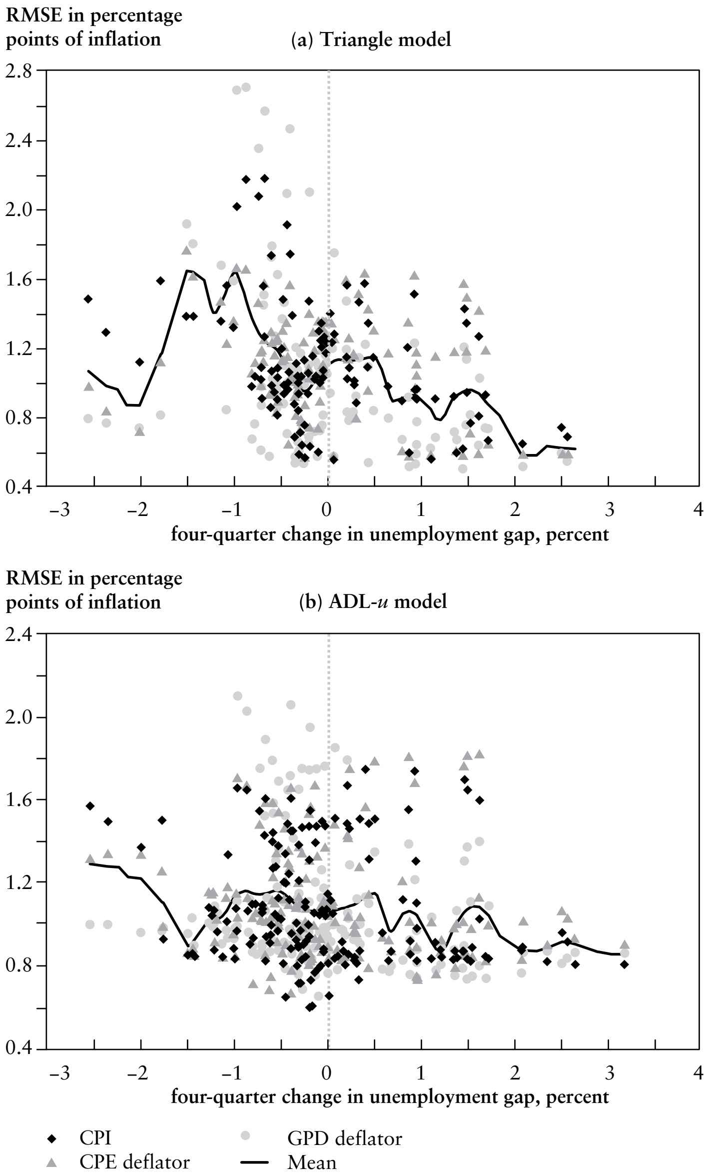


Figure 3.14
Scatterplot of Rolling Root Mean Square Errors of Headline Inflation Forecasts, Relative to Unobserved Components-Stochastic Volatility, vs. the Four-quarter Change in the Unemployment Gap (two-sided bandpass)
Note: Mean is kernel regression estimate using data for all three series. Each point represents a quarter.

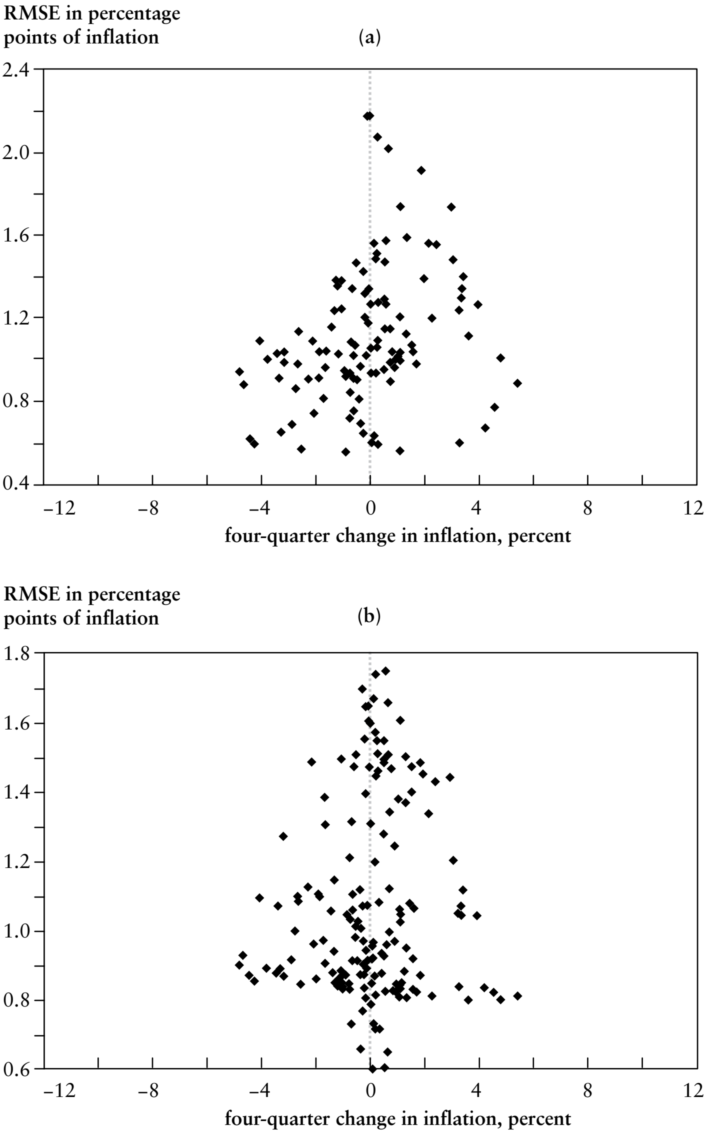


Figure 3.15
Scatterplot of Rolling Root Mean Square Errors of CPI-All Inflation Forecasts from (a) Triangle Model and (b) ADL-*u* Model, Relative to Unobserved Components-Stochastic Volatility Model, versus the 4-quarter Change in 4-quarter Inflation. Each point represents a quarter.

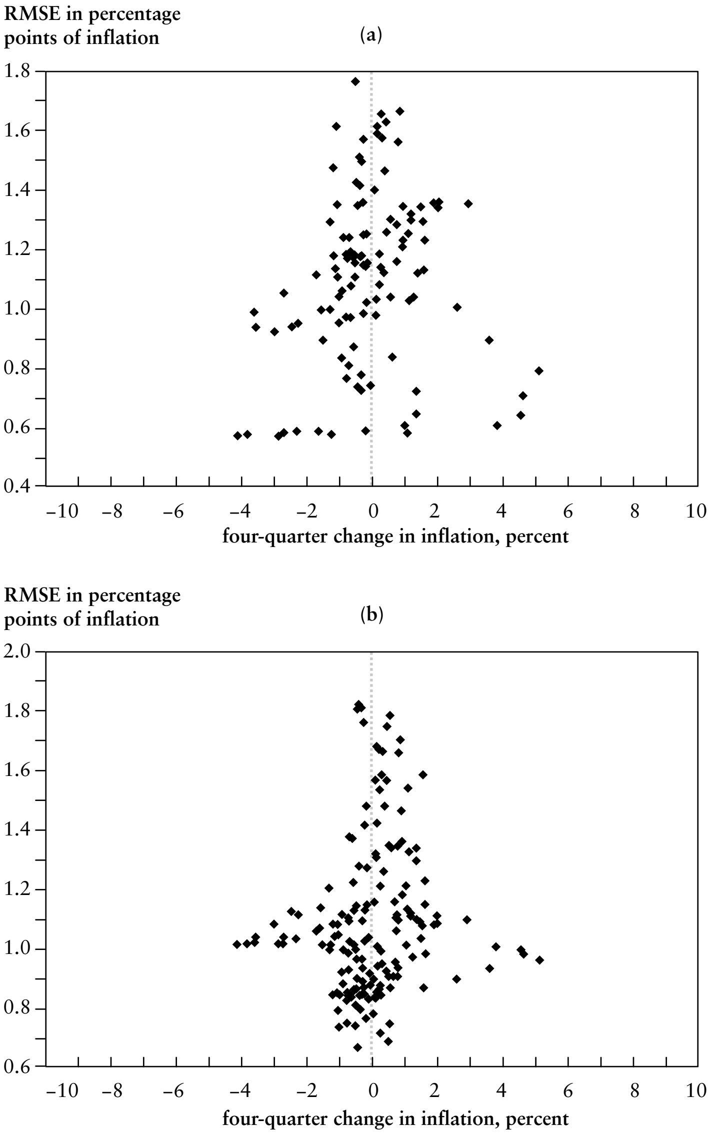


Figure 3.16
Scatterplot of Rolling Root Mean Square Errors of PCE-All Inflation Forecasts from (a) Triangle Model and (b) ADL-*u* Model, Relative to Unobserved Components-Stochastic Volatility Model, versus the 4-quarter Change in 4-quarter Inflation. Each point represents a quarter.

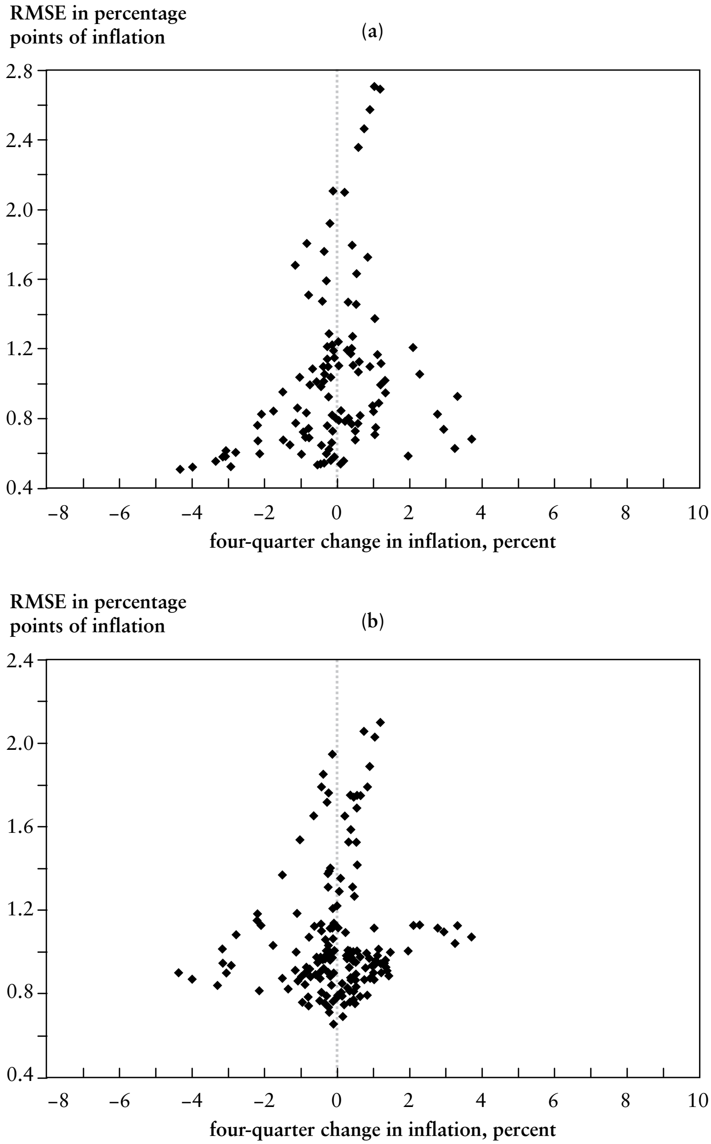


Figure 3.17
Scatterplot of Rolling Root Mean Square Errors of GDP Inflation Forecasts from (a) Triangle Model and (b) ADL-*u* Model, Relative to Unobserved Components-Stochastic Volatility Model, versus the 4-quarter Change in 4-quarter Inflation. Each point represents a quarter.

triangle model applied to the GDP deflator, the episodes of best performance are not associated with large changes in inflation.

As presented here, these patterns cannot yet be used to improve forecasts: the sharpest patterns are ones that appear using two-sided gaps. Still, these results point to a possible resolution of the Atkeson-Ohanian conundrum in which real economic activity seems to play no role in inflation forecasting. The results here suggest that, if times are quiet—if the unemployment rate is close to the NAIRU—then in fact one is better off using a univariate forecast than introducing additional estimation error by making a multivariate forecast. But if the unemployment rate is far from the NAIRU, then knowledge of that large unemployment gap is useful for inflation forecasting.

■ *We thank Ian Dew-Becker for research assistance, Michelle Barnes of the Federal Reserve Bank of Boston for data assistance, and an anonymous referee for some important references. This research was funded in part by National Science Foundation grant SBR-0617811. Data and replication files are available at http://www.princeton.edu/~mwatson.*

Table 3.1
Root Mean Squared Errors for Inflation Forecasting Models by Subperiod, Relative to the Unobserved Components-Stochastic Volatility Model: CPI-all

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
Number of observations	32	36	32	32	32	25
Root MSE of UC-SV forecast	0.82	1.99	2.35	1.39	0.68	1.05
Forecasting model and relative root mean square errors						
Univariate forecasts						
UC-SV	1.00	1.00	1.00	1.00	1.00	1.00
AR(AIC)_rec	.	1.09	1.05	1.12	1.03	1.39
AR(AIC)_iter_rec	.	1.06	1.00	1.12	1.02	1.43
AR(BIC)_rec	.	1.10	1.03	1.10	1.03	1.37
AO	1.01	1.23	1.12	1.00	1.10	1.14
MA(1)_rec	.	1.07	1.01	1.07	1.03	1.37
AR(4)_rec	.	1.12	1.02	1.13	1.02	1.42
AR(AIC)_roll	.	1.10	1.09	1.03	1.21	1.30
AR(AIC)_iter_roll	.	1.08	1.03	1.15	1.11	1.37
AR(BIC)_roll	.	1.09	1.08	1.02	1.14	1.32
AR(4)_roll	.	1.19	1.06	1.07	1.17	1.29
AR(24)_iter	.	.	.	1.30	1.04	1.33
AR(24)_iter_nocon	.	.	1.18	1.25	1.00	1.32
MA(1)_roll	.	1.04	1.02	1.07	1.05	1.13
MA(2) - NS	0.98	1.14	1.13	0.95	1.01	1.12
MA(1), θ=.25	1.12	1.01	1.00	1.11	1.06	1.52
MA(1), θ=.65	0.97	1.15	1.12	0.96	1.03	1.12
Single-predictor ADL forecasts						
UR(Level)_AIC_rec	.	0.96	0.92	0.98	1.28	1.36
UR(Dif)_AIC_rec	.	0.93	0.94	1.04	1.22	1.39
UR(1sdBP)_AIC_rec	.	0.96	0.95	1.00	1.22	1.38
GDP(Dif)_AIC_rec	.	0.88	0.93	1.00	1.09	1.36
GDP(1sdBP)_AIC_rec	.	1.03	0.90	1.00	1.08	1.34
IP(Dif)_AIC_rec	.	0.89	0.93	1.02	1.22	1.43
IP(1sdBP)_AIC_rec	.	0.95	0.93	1.01	1.17	1.40
Emp(Dif)_AIC_rec	.	0.93	0.86	1.01	1.06	1.53
Emp(1sdBP)_AIC_rec	.	0.95	0.87	1.02	1.14	1.49
CapU(Level)_AIC_rec	.	.	.	1.03	1.39	1.56
CapU((Dif)_AIC_rec	.	.	.	1.03	1.30	1.45
CapU(1sdBP)_AIC_rec	.	.	.	0.99	1.21	1.35
HPerm(Level)_AIC_rec	.	.	0.79	1.12	1.14	1.75
HPerm((Dif)_AIC_rec	.	.	0.91	1.29	0.97	1.67
HPerm(1sdBP)_AIC_rec	.	.	0.90	1.02	1.08	1.37
CFNAI(Dif)_AIC_rec	.	.	.	1.01	1.21	1.57
CFNAI(1sdBP)_AIC_rec	.	.	.	0.98	1.18	1.42
UR_5wk(Level)_AIC_rec	.	1.06	0.93	1.05	1.73	1.38
UR_5wk(Dif)_AIC_rec	.	0.94	0.91	1.07	1.34	1.40
UR_5wk(1sdBP)_AIC_rec	.	0.97	0.90	1.06	1.34	1.31
AHE(Dif)_AIC_rec	.	.	1.10	1.19	1.03	1.48

Table 3.1 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
AHE(1sdBP)_AIC_rec	.	.	1.12	1.20	1.01	1.46
RealAHE(Dif)_AIC_rec	.	.	1.10	1.19	1.03	1.48
RealAHE(1sdBP)_AIC_rec	.	.	1.12	1.20	1.01	1.46
LaborShare(Level)_AIC_rec	.	1.06	1.02	1.21	1.76	1.44
LaborShare(Dif)_AIC_rec	.	1.08	1.03	1.12	1.06	1.36
ULaborShare(1sdBP)_AIC_rec	.	1.10	1.01	1.09	1.30	1.36
CPI_Med(Level)_AIC_rec	.	.	.	1.34	1.39	1.54
CPI_Med(Dif)_AIC_rec	.	.	.	1.20	1.11	1.45
CPI_TrMn(Level)_AIC_rec	.	.	.	1.35	1.46	1.47
CPI_TrMn(Dif)_AIC_rec	.	.	.	1.10	1.07	1.45
ExRate(Dif)_AIC_rec	.	.	.	1.43	1.26	1.21
ExRate(1sdBP)_AIC_rec	.	.	.	1.82	1.04	1.28
tb_spr_AIC_rec	.	1.10	1.05	1.21	1.24	1.56
UR(Level)_AIC_roll	.	1.20	1.13	0.99	1.32	1.30
UR(Dif)_AIC_roll	.	1.07	1.00	1.04	1.23	1.28
UR(1sdBP)_AIC_roll	.	1.17	1.07	1.03	1.28	1.30
GDP(Dif)_AIC_roll	.	1.01	1.01	0.98	1.36	1.25
GDP(1sdBP)_AIC_roll	.	1.10	0.91	1.00	1.25	1.25
IP(Dif)_AIC_roll	.	0.95	0.99	1.05	1.26	1.33
IP(1sdBP)_AIC_roll	.	1.07	1.00	1.05	1.30	1.28
Emp(Dif)_AIC_roll	.	1.06	0.97	0.99	1.23	1.24
Emp(1sdBP)_AIC_roll	.	1.19	0.91	1.02	1.26	1.31
CapU(Level)_AIC_roll	.	.	.	0.98	1.38	1.33
CapU((Dif)_AIC_roll	.	.	.	1.02	1.27	1.29
CapU(1sdBP)_AIC_roll	.	.	.	0.97	1.35	1.22
HPerm(Level)_AIC_roll	.	.	0.75	1.27	1.23	1.41
HPerm((Dif)_AIC_roll	.	.	1.14	1.16	1.05	1.55
HPerm(1sdBP)_AIC_roll	.	.	0.94	1.21	1.20	1.32
CFNAI(Dif)_AIC_roll	.	.	.	0.97	1.28	1.25
CFNAI(1sdBP)_AIC_roll	.	.	.	1.02	1.28	1.25
UR_5wk(Level)_AIC_roll	.	1.19	0.93	1.18	1.60	1.34
UR_5wk(Dif)_AIC_roll	.	1.03	0.97	1.03	1.41	1.27
UR_5wk(1sdBP)_AIC_roll	.	1.08	0.85	1.06	1.45	1.31
AHE(Dif)_AIC_roll	.	.	1.10	1.08	1.38	1.24
AHE(1sdBP)_AIC_roll	.	.	1.12	1.07	1.33	1.19
RealAHE(Dif)_AIC_roll	.	.	1.10	1.08	1.38	1.24
RealAHE(1sdBP)_AIC_roll	.	.	1.12	1.07	1.33	1.19
LaborShare(Level)_AIC_roll	.	1.15	1.02	1.12	1.31	1.32
LaborShare(Dif)_AIC_roll	.	1.13	1.09	1.02	1.63	1.32
ULaborShare(1sdBP)_AIC_roll	.	1.18	1.09	1.04	1.31	1.30
CPI_Med(Level)_AIC_roll	.	.	.	1.15	1.34	1.18
CPI_Med(Dif)_AIC_roll	.	.	.	1.01	1.15	1.29
CPI_TrMn(Level)_AIC_roll	.	.	.	1.12	1.38	1.28
CPI_TrMn(Dif)_AIC_roll	.	.	.	1.05	1.15	1.31
ExRate(Dif)_AIC_roll	.	.	.	1.53	1.20	1.28
ExRate(1sdBP)_AIC_roll	.	.	.	1.91	1.16	1.34

Table 3.1 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
tb_spr_AIC_roll	.	1.13	1.23	1.33	1.42	1.37
UR(Level)_BIC_rec	.	0.92	0.91	0.98	1.28	1.36
UR(Dif)_BIC_rec	.	0.88	0.94	1.06	1.16	1.35
UR(1sdBP)_BIC_rec	.	0.91	0.93	0.96	1.17	1.35
GDP(Dif)_BIC_rec	.	0.95	0.99	1.00	1.09	1.36
GDP(1sdBP)_BIC_rec	.	0.99	0.95	0.99	1.05	1.33
IP(Dif)_BIC_rec	.	0.90	0.97	1.03	1.20	1.41
IP(1sdBP)_BIC_rec	.	0.95	0.99	1.00	1.11	1.39
Emp(Dif)_BIC_rec	.	0.90	0.92	0.98	1.05	1.51
Emp(1sdBP)_BIC_rec	.	0.93	0.93	0.99	1.09	1.45
CapU(Level)_BIC_rec	.	.	.	1.02	1.29	1.56
CapU((Dif)_BIC_rec	.	.	.	1.07	1.30	1.46
CapU(1sdBP)_BIC_rec	.	.	.	0.97	1.17	1.30
HPerm(Level)_BIC_rec	.	.	0.82	1.06	1.14	1.75
HPerm((Dif)_BIC_rec	.	.	1.05	1.32	0.97	1.65
HPerm(1sdBP)_BIC_rec	.	.	0.93	1.02	1.08	1.37
CFNAI(Dif)_BIC_rec	.	.	.	0.92	1.18	1.44
CFNAI(1sdBP)_BIC_rec	.	.	.	0.95	1.18	1.42
UR_5wk(Level)_BIC_rec	.	1.03	0.92	1.13	1.62	1.48
UR_5wk(Dif)_BIC_rec	.	0.94	0.96	1.15	1.17	1.49
UR_5wk(1sdBP)_BIC_rec	.	0.94	0.88	1.11	1.27	1.34
AHE(Dif)_BIC_rec	.	.	1.08	1.19	1.10	1.42
AHE(1sdBP)_BIC_rec	.	.	1.10	1.23	1.05	1.37
RealAHE(Dif)_BIC_rec	.	.	1.08	1.19	1.10	1.42
RealAHE(1sdBP)_BIC_rec	.	.	1.10	1.23	1.05	1.37
LaborShare(Level)_BIC_rec	.	1.02	0.99	1.20	1.61	1.44
LaborShare(Dif)_BIC_rec	.	1.08	1.03	1.13	1.07	1.36
ULaborShare(1sdBP)_BIC_rec	.	1.07	0.97	1.13	1.30	1.40
CPI_Med(Level)_BIC_rec	.	.	.	1.22	1.44	1.53
CPI_Med(Dif)_BIC_rec	.	.	.	1.21	1.14	1.51
CPI_TrMn(Level)_BIC_rec	.	.	.	1.23	1.43	1.49
CPI_TrMn(Dif)_BIC_rec	.	.	.	1.10	1.07	1.49
ExRate(Dif)_BIC_rec	.	.	.	1.53	1.19	1.32
ExRate(1sdBP)_BIC_rec	.	.	.	1.87	1.09	1.28
tb_spr_BIC_rec	.	1.09	1.09	1.17	1.06	1.40
UR(Level)_BIC_roll	.	1.16	1.05	0.99	1.33	1.31
UR(Dif)_BIC_roll	.	0.99	0.99	0.99	1.19	1.35
UR(1sdBP)_BIC_roll	.	1.14	0.98	0.96	1.24	1.28
GDP(Dif)_BIC_roll	.	0.96	1.01	0.98	1.28	1.31
GDP(1sdBP)_BIC_roll	.	1.04	0.92	0.98	1.18	1.30
IP(Dif)_BIC_roll	.	0.99	1.01	1.05	1.24	1.29
IP(1sdBP)_BIC_roll	.	1.08	0.96	0.97	1.31	1.32
Emp(Dif)_BIC_roll	.	1.06	0.95	1.02	1.22	1.27
Emp(1sdBP)_BIC_roll	.	1.12	0.92	1.05	1.24	1.27
CapU(Level)_BIC_roll	.	.	.	0.97	1.30	1.30
CapU(Dif)_BIC_roll	.	.	.	1.01	1.26	1.27

Table 3.1 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
CapU(1sdBP)_BIC_roll	.	.	.	0.93	1.23	1.25
HPerm(Level)_BIC_roll	.	.	0.77	1.25	1.22	1.43
HPerm((Dif)_BIC_roll	.	.	1.09	1.21	1.06	1.36
HPerm(1sdBP)_BIC_roll	.	.	0.92	1.21	1.19	1.33
CFNAI(Dif)_BIC_roll	.	.	.	0.96	1.26	1.32
CFNAI(1sdBP)_BIC_roll	.	.	.	1.02	1.22	1.24
UR_5wk(Level)_BIC_roll	.	1.18	0.95	1.19	1.35	1.43
UR_5wk(Dif)_BIC_roll	.	0.96	0.99	1.10	1.19	1.35
UR_5wk(1sdBP)_BIC_roll	.	1.04	0.93	1.09	1.32	1.37
AHE(Dif)_BIC_roll	.	.	1.09	1.03	1.19	1.38
AHE(1sdBP)_BIC_roll	.	.	1.10	1.16	1.12	1.23
RealAHE(Dif)_BIC_roll	.	.	1.09	1.03	1.19	1.38
RealAHE(1sdBP)_BIC_roll	.	.	1.10	1.16	1.12	1.23
LaborShare(Level)_BIC_roll	.	1.09	1.05	1.05	1.23	1.33
LaborShare(Dif)_BIC_roll	.	1.12	1.10	1.11	1.17	1.38
ULaborShare(1sdBP)_BIC_roll	.	1.11	1.06	1.02	1.29	1.36
CPI_Med(Level)_BIC_roll	.	.	.	1.07	1.28	1.24
CPI_Med(Dif)_BIC_roll	.	.	.	1.04	1.15	1.30
CPI_TrMn(Level)_BIC_roll	.	.	.	1.02	1.28	1.30
CPI_TrMn(Dif)_BIC_roll	.	.	.	1.07	1.14	1.36
ExRate(Dif)_BIC_roll	.	.	.	1.56	1.13	1.34
ExRate(1sdBP)_BIC_roll	.	.	.	1.95	1.11	1.36
tb_spr_BIC_roll	.	1.16	1.25	1.32	1.22	1.38
Triangle model forecasts						
Triangle Constant NAIRU	.	.	0.94	1.11	1.14	1.11
Triangle TV NAIRU	.	.	0.95	1.15	1.07	1.16
Triangle Constant NAIRU (no z)	.	.	1.02	1.19	1.34	1.34
Triangle TV NAIRU (no z)	.	.	1.12	1.23	1.10	1.52
Combination forecasts						
Activity Median Combining	.	0.96	0.88	0.96	1.13	1.30
Activity Mean Combining	.	0.97	0.86	0.96	1.11	1.30
Activity Tr. Mean Combining	.	0.97	0.87	0.96	1.11	1.30
Activity MSE(A) Combining	.	.	0.86	0.97	1.12	1.31
Activity MSE(B) Combining	.	.	0.86	0.96	1.12	1.31
Activity MSE(C) Combining	.	.	0.86	0.96	1.11	1.30
Activity MSE(D) Combining	.	.	0.86	0.98	1.14	1.33
Activity MSE(E) Combining	.	.	0.87	0.97	1.13	1.32
Activity MSE(F) Combining	.	.	0.87	0.96	1.12	1.30
Activity Rec. Best(4q) Combining	.	1.12	0.74	0.99	1.38	1.56
Activity Rec. Best(8q) Combining	.	1.07	0.90	1.22	1.48	1.36
OtherADL Median Combining	.	1.07	1.06	1.03	1.11	1.29
OtherADL Mean Combining	.	1.08	1.01	1.06	1.09	1.30
OtherADL Tr. Mean Combining	.	1.08	1.03	1.05	1.09	1.31
OtherADL MSE(A) Combining	.	.	0.98	1.07	1.11	1.30
OtherADL MSE(B) Combining	.	.	0.98	1.07	1.12	1.30
OtherADL MSE(C) Combining	.	.	0.99	1.07	1.12	1.30

Table 3.1 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
OtherADL MSE(D) Combining	.	.	0.98	1.08	1.13	1.31
OtherADL MSE(E) Combining	.	.	0.99	1.07	1.13	1.30
OtherADL MSE(F) Combining	.	.	0.99	1.07	1.14	1.30
OtherADL Rec. Best(4q) Combining	.	1.13	1.05	1.12	1.36	1.37
OtherADL Rec. Best(8q) Combining	.	1.14	1.09	1.21	1.30	1.42
All Median Combining	.	0.98	0.92	0.98	1.10	1.31
All Mean Combining	.	0.99	0.89	0.98	1.07	1.29
All Tr. Mean Combining	.	0.99	0.90	0.98	1.08	1.30
All MSE(A) Combining	.	.	0.87	0.99	1.10	1.30
All MSE(B) Combining	.	.	0.87	0.98	1.10	1.30
All MSE(C) Combining	.	.	0.87	0.98	1.09	1.29
All MSE(D) Combining	.	.	0.87	1.00	1.12	1.31
All MSE(E) Combining	.	.	0.88	0.99	1.12	1.30
All MSE(F) Combining	.	.	0.88	0.98	1.10	1.29
All Rec. Best(4q) Combining	.	1.12	0.74	1.11	1.47	1.63
All Rec. Best(8q) Combining	.	1.08	0.92	1.19	1.51	1.43
UCSV and Triangle Rec. Best(4q) Combining	.	.	.	1.02	1.05	1.01
UCSV and Triangle Rec. Best(8q) Combining	.	.	.	1.06	1.05	1.11

Notes to Table 3.1: Entries are Root Mean Squared Errors, relative to the Root Mean Squared Errors of the Unobserved Components-Stochastic Volatility model, over the indicated sample period. Blanks indicate insufficient data to compute forecasts over the indicated subsample. The abbreviations denote:

_AIC: AIC lag selection, up to six lags (for ADL models, AIC over the two lag lengths separately)
_BIC: BIC lag selection, up to six lags (for ADL models, AIC over the two lag lengths separately)
_rec: recursive estimation
_roll: rolling estimation
Level: indicated predictor appears in levels
Dif: indicated predictor appears in log differences
1sdBP: indicated predictor appears in gap form, computed using 1-sided bandpass filter as discussed in the text
Triangle: Triangle model or TV-triangle model, with or without supply shock ("z") variables
mean, median, trimmed mean: forecast combining methods, for the indicated group of forecasts
MSE(A) – MSE(F): MSE-based combining as indicated in equations (15)–(20).
Best (four-quarter) and Best (eight-quarter): recently best forecast based on cumulative MSE over past four (or eight) quarters
UCSV and Triangle Rec. Best (four-quarter) and (eight-quarter) Combining: best of UC-SV and triangle models (constant NAIRU) based on cumulative MSE over past four (or eight) quarters

nocon: constant term is suppressed

Source: Authors' calculations.

Table 3.2
Root Mean Squared Errors for Inflation Forecasting Models by Subperiod, Relative to the Unobserved Components-Stochastic Volatility Model: CPI-core

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
Number of observations	32	36	32	32	32	25
Root MSE of UC-SV forecast	0.82	2.15	2.30	0.58	0.31	0.53
Forecasting model and relative root mean square errors						
Univariate forecasts						
UC-SV	1.00	1.00	1.00	1.00	1.00	1.00
AR(AIC)_rec	.	.	1.07	1.07	1.04	1.05
AR(AIC)_iter_rec	.	.	1.08	1.06	1.04	1.06
AR(BIC)_rec	.	.	1.07	1.01	1.06	1.05
AO	1.03	1.14	1.01	1.08	1.04	1.06
MA(1)_rec	.	1.05	1.04	1.00	1.01	1.04
AR(4)_rec	.	.	1.09	1.09	1.03	1.04
AR(AIC)_roll	.	.	1.12	1.05	1.12	1.09
AR(AIC)_iter_roll	.	.	1.21	1.15	1.15	1.11
AR(BIC)_roll	.	.	1.11	1.03	1.11	1.09
AR(4)_roll	.	.	1.15	1.06	1.12	1.10
AR(24)_iter	.	.	1.24	1.57	1.51	0.93
AR(24)_iter_nocon	.	.	1.10	1.23	1.32	0.91
MA(1)_roll	.	1.04	1.03	1.12	1.00	1.07
MA(2) - NS	1.04	1.04	1.01	1.04	1.07	0.98
MA(1), θ=.25	1.05	1.01	0.98	1.02	1.04	1.07
MA(1), θ=.65	1.03	1.06	1.00	1.03	1.04	1.00
Single-predictor ADL forecasts						
UR(Level)_AIC_rec	.	.	0.89	0.83	1.92	1.11
UR(Dif)_AIC_rec	.	.	0.95	0.91	1.43	1.01
UR(1sdBP)_AIC_rec	.	.	0.91	1.01	1.63	1.05
GDP(Dif)_AIC_rec	.	.	1.02	0.91	1.02	0.92
GDP(1sdBP)_AIC_rec	.	.	0.95	1.00	1.17	1.09
IP(Dif)_AIC_rec	.	.	1.03	1.00	1.25	1.16
IP(1sdBP)_AIC_rec	.	.	0.97	0.85	1.54	1.39
Emp(Dif)_AIC_rec	.	.	0.92	0.90	1.18	1.32
Emp(1sdBP)_AIC_rec	.	.	0.91	0.90	1.28	1.27
CapU(Level)_AIC_rec	.	.	.	1.24	2.00	2.19
CapU((Dif)_AIC_rec	.	.	.	1.12	1.21	1.25
CapU(1sdBP)_AIC_rec	.	.	.	1.34	1.26	1.20
HPerm(Level)_AIC_rec	.	.	0.91	1.29	1.46	1.91
HPerm((Dif)_AIC_rec	.	.	1.06	1.10	1.21	1.04
HPerm(1sdBP)_AIC_rec	.	.	0.99	1.07	1.48	0.98
CFNAI(Dif)_AIC_rec	.	.	.	1.16	1.27	1.39
CFNAI(1sdBP)_AIC_rec	.	.	.	1.17	1.29	1.37
UR_5wk(Level)_AIC_rec	.	.	0.86	1.10	3.09	1.32
UR_5wk(Dif)_AIC_rec	.	.	0.91	1.19	1.91	1.08
UR_5wk(1sdBP)_AIC_rec	.	.	0.90	1.22	2.32	1.06
AHE(Dif)_AIC_rec	.	.	1.12	1.08	1.03	1.06

Table 3.2 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
AHE(1sdBP)_AIC_rec	.	.	1.16	1.36	1.10	1.10
RealAHE(Dif)_AIC_rec	.	.	1.12	1.08	1.03	1.06
RealAHE(1sdBP)_AIC_rec	.	.	1.16	1.36	1.10	1.10
LaborShare(Level)_AIC_rec	.	.	1.11	1.37	2.18	1.12
LaborShare(Dif)_AIC_rec	.	.	1.14	1.05	1.27	1.06
ULaborShare(1sdBP)_AIC_rec	.	.	1.14	1.15	1.58	1.07
CPI_Med(Level)_AIC_rec	.	.	.	1.30	2.06	1.10
CPI_Med(Dif)_AIC_rec	.	.	.	1.14	1.81	1.25
CPI_TrMn(Level)_AIC_rec	.	.	.	1.28	1.69	1.23
CPI_TrMn(Dif)_AIC_rec	.	.	.	1.11	1.38	1.20
ExRate(Dif)_AIC_rec	.	.	.	2.97	1.43	0.93
ExRate(1sdBP)_AIC_rec	.	.	.	3.14	1.25	1.24
tb_spr_AIC_rec	.	.	1.10	1.52	2.53	1.32
UR(Level)_AIC_roll	.	.	1.27	1.52	1.34	1.19
UR(Dif)_AIC_roll	.	.	1.02	1.26	1.07	1.07
UR(1sdBP)_AIC_roll	.	.	0.87	1.40	1.12	1.21
GDP(Dif)_AIC_roll	.	.	1.12	1.37	1.12	1.10
GDP(1sdBP)_AIC_roll	.	.	1.00	1.52	1.02	1.10
IP(Dif)_AIC_roll	.	.	1.08	1.48	1.15	1.26
IP(1sdBP)_AIC_roll	.	.	0.89	1.70	1.14	1.39
Emp(Dif)_AIC_roll	.	.	0.94	1.47	1.09	1.31
Emp(1sdBP)_AIC_roll	.	.	0.84	1.57	1.08	1.54
CapU(Level)_AIC_roll	.	.	.	1.59	1.31	1.39
CapU((Dif)_AIC_roll	.	.	.	1.48	1.11	1.17
CapU(1sdBP)_AIC_roll	.	.	.	1.48	1.26	1.26
HPerm(Level)_AIC_roll	.	.	0.89	2.35	1.04	1.24
HPerm((Dif)_AIC_roll	.	.	1.10	1.63	1.12	1.14
HPerm(1sdBP)_AIC_roll	.	.	1.04	1.95	1.05	1.12
CFNAI(Dif)_AIC_roll	.	.	.	1.44	1.03	1.17
CFNAI(1sdBP)_AIC_roll	.	.	.	1.50	0.92	1.25
UR_5wk(Level)_AIC_roll	.	.	1.05	2.12	1.28	1.13
UR_5wk(Dif)_AIC_roll	.	.	1.09	1.44	1.08	1.10
UR_5wk(1sdBP)_AIC_roll	.	.	0.85	1.32	1.33	1.12
AHE(Dif)_AIC_roll	.	.	1.23	1.23	1.13	1.16
AHE(1sdBP)_AIC_roll	.	.	1.22	1.53	1.12	1.15
RealAHE(Dif)_AIC_roll	.	.	1.23	1.23	1.13	1.16
RealAHE(1sdBP)_AIC_roll	.	.	1.22	1.53	1.12	1.15
LaborShare(Level)_AIC_roll	.	.	1.20	1.34	1.83	1.12
LaborShare(Dif)_AIC_roll	.	.	1.41	1.15	1.97	1.10
ULaborShare(1sdBP)_AIC_roll	.	.	1.30	1.16	1.69	1.10
CPI_Med(Level)_AIC_roll	.	.	.	1.37	1.69	0.80
CPI_Med(Dif)_AIC_roll	.	.	.	1.04	1.25	1.15
CPI_TrMn(Level)_AIC_roll	.	.	.	1.30	1.53	1.19
CPI_TrMn(Dif)_AIC_roll	.	.	.	1.08	1.14	1.19
ExRate(Dif)_AIC_roll	.	.	.	3.48	1.19	1.11
ExRate(1sdBP)_AIC_roll	.	.	.	3.58	1.05	1.10

Table 3.2 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
tb_spr_AIC_roll	.	.	1.15	1.61	1.22	1.12
UR(Level)_BIC_rec	.	.	0.97	0.83	1.92	1.11
UR(Dif)_BIC_rec	.	.	1.00	0.91	1.43	1.01
UR(1sdBP)_BIC_rec	.	.	0.88	1.01	1.62	1.05
GDP(Dif)_BIC_rec	.	.	1.04	0.98	1.06	0.91
GDP(1sdBP)_BIC_rec	.	.	0.96	0.96	1.14	1.07
IP(Dif)_BIC_rec	.	.	1.04	0.99	1.11	1.11
IP(1sdBP)_BIC_rec	.	.	1.00	0.90	1.41	1.34
Emp(Dif)_BIC_rec	.	.	0.97	0.90	1.18	1.32
Emp(1sdBP)_BIC_rec	.	.	0.83	0.90	1.28	1.27
CapU(Level)_BIC_rec	.	.	.	1.24	1.83	2.10
CapU((Dif)_BIC_rec	.	.	.	1.06	1.11	1.22
CapU(1sdBP)_BIC_rec	.	.	.	1.22	1.23	1.27
HPerm(Level)_BIC_rec	.	.	0.91	1.27	1.46	1.91
HPerm((Dif)_BIC_rec	.	.	1.05	1.16	1.21	1.04
HPerm(1sdBP)_BIC_rec	.	.	0.99	1.07	1.48	0.98
CFNAI(Dif)_BIC_rec	.	.	.	1.10	1.27	1.29
CFNAI(1sdBP)_BIC_rec	.	.	.	1.17	1.29	1.37
UR_5wk(Level)_BIC_rec	.	.	0.93	1.09	2.88	1.42
UR_5wk(Dif)_BIC_rec	.	.	1.01	1.19	1.94	1.15
UR_5wk(1sdBP)_BIC_rec	.	.	0.87	1.24	2.23	1.05
AHE(Dif)_BIC_rec	.	.	1.13	1.03	1.05	1.09
AHE(1sdBP)_BIC_rec	.	.	1.17	1.24	1.16	1.09
RealAHE(Dif)_BIC_rec	.	.	1.13	1.03	1.05	1.09
RealAHE(1sdBP)_BIC_rec	.	.	1.17	1.24	1.16	1.09
LaborShare(Level)_BIC_rec	.	.	1.08	1.22	1.68	1.09
LaborShare(Dif)_BIC_rec	.	.	1.09	1.02	1.34	1.08
ULaborShare(1sdBP)_BIC_rec	.	.	1.09	1.08	1.27	1.02
CPI_Med(Level)_BIC_rec	.	.	.	1.34	1.66	1.11
CPI_Med(Dif)_BIC_rec	.	.	.	1.02	1.25	1.10
CPI_TrMn(Level)_BIC_rec	.	.	.	1.28	1.71	1.23
CPI_TrMn(Dif)_BIC_rec	.	.	.	1.09	1.38	1.20
ExRate(Dif)_BIC_rec	.	.	.	1.73	1.34	1.09
ExRate(1sdBP)_BIC_rec	.	.	.	2.85	1.10	1.01
tb_spr_BIC_rec	.	.	1.07	1.50	2.40	1.32
UR(Level)_BIC_roll	.	.	1.16	1.49	1.26	1.16
UR(Dif)_BIC_roll	.	.	1.04	1.18	1.06	1.08
UR(1sdBP)_BIC_roll	.	.	0.88	1.33	1.15	1.10
GDP(Dif)_BIC_roll	.	.	1.01	1.30	1.11	1.04
GDP(1sdBP)_BIC_roll	.	.	0.95	1.47	1.10	1.11
IP(Dif)_BIC_roll	.	.	1.06	1.49	1.15	1.23
IP(1sdBP)_BIC_roll	.	.	0.86	1.67	1.13	1.31
Emp(Dif)_BIC_roll	.	.	1.00	1.45	1.15	1.16
Emp(1sdBP)_BIC_roll	.	.	0.84	1.56	1.15	1.50
CapU(Level)_BIC_roll	.	.	.	1.61	1.32	1.37
CapU(Dif)_BIC_roll	.	.	.	1.48	1.11	1.15

Table 3.2 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
CapU(1sdBP)_BIC_roll	.	.	.	1.46	1.30	1.29
HPerm(Level)_BIC_roll	.	.	0.89	2.29	1.05	1.17
HPerm((Dif)_BIC_roll	.	.	1.10	1.62	1.20	1.13
HPerm(1sdBP)_BIC_roll	.	.	1.05	1.93	1.08	1.11
CFNAI(Dif)_BIC_roll	.	.	.	1.47	1.03	1.13
CFNAI(1sdBP)_BIC_roll	.	.	.	1.53	0.91	1.20
UR_5wk(Level)_BIC_roll	.	.	1.08	1.96	1.19	1.12
UR_5wk(Dif)_BIC_roll	.	.	1.11	1.49	1.11	1.09
UR_5wk(1sdBP)_BIC_roll	.	.	0.85	1.36	1.31	1.08
AHE(Dif)_BIC_roll	.	.	1.12	1.15	1.12	1.13
AHE(1sdBP)_BIC_roll	.	.	1.16	1.32	1.13	1.16
RealAHE(Dif)_BIC_roll	.	.	1.12	1.15	1.12	1.13
RealAHE(1sdBP)_BIC_roll	.	.	1.16	1.32	1.13	1.16
LaborShare(Level)_BIC_roll	.	.	1.10	1.35	1.61	1.11
LaborShare(Dif)_BIC_roll	.	.	1.20	1.21	1.84	1.10
ULaborShare(1sdBP)_BIC_roll	.	.	1.13	1.12	1.56	1.10
CPI_Med(Level)_BIC_roll	.	.	.	1.41	1.47	0.84
CPI_Med(Dif)_BIC_roll	.	.	.	1.05	1.22	1.18
CPI_TrMn(Level)_BIC_roll	.	.	.	1.23	1.47	1.19
CPI_TrMn(Dif)_BIC_roll	.	.	.	1.08	1.15	1.18
ExRate(Dif)_BIC_roll	.	.	.	3.03	1.10	1.14
ExRate(1sdBP)_BIC_roll	.	.	.	3.35	1.06	1.06
tb_spr_BIC_roll	.	.	1.14	1.45	1.28	1.11
Triangle model forecasts						
Triangle Constant NAIRU	.	.	1.32	1.50	1.81	1.44
Triangle TV NAIRU	.	.	1.32	1.46	1.48	1.39
Triangle Constant NAIRU (no z)	.	.	1.05	1.11	2.34	1.17
Triangle TV NAIRU (no z)	.	.	1.07	1.22	1.63	1.23
Combination forecasts						
Activity Median Combining	.	.	0.86	0.86	1.00	1.07
Activity Mean Combining	.	.	0.86	0.89	1.02	1.02
Activity Tr. Mean Combining	.	.	0.86	0.88	1.01	1.04
Activity MSE(A) Combining	.	.	.	0.87	1.05	1.05
Activity MSE(B) Combining	.	.	.	0.88	1.06	1.04
Activity MSE(C) Combining	.	.	.	0.88	1.06	1.04
Activity MSE(D) Combining	.	.	.	0.87	1.09	1.07
Activity MSE(E) Combining	.	.	.	0.87	1.10	1.06
Activity MSE(F) Combining	.	.	.	0.88	1.11	1.05
Activity Rec. Best(4q) Combining	.	.	1.13	1.15	1.22	1.19
Activity Rec. Best(8q) Combining	.	.	0.96	1.40	1.50	1.40
OtherADL Median Combining	.	.	1.08	0.99	1.11	1.06
OtherADL Mean Combining	.	.	1.05	1.04	1.15	1.02
OtherADL Tr. Mean Combining	.	.	1.07	0.96	1.12	1.03
OtherADL MSE(A) Combining	.	.	.	1.02	1.16	1.04
OtherADL MSE(B) Combining	.	.	.	1.00	1.18	1.03
OtherADL MSE(C) Combining	.	.	.	0.99	1.18	1.03

Table 3.2 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
OtherADL MSE(D) Combining	.	.	.	1.07	1.16	1.06
OtherADL MSE(E) Combining	.	.	.	1.02	1.17	1.05
OtherADL MSE(F) Combining	.	.	.	1.00	1.18	1.04
OtherADL Rec. Best(4q) Combining	.	.	1.08	2.06	1.18	1.06
OtherADL Rec. Best(8q) Combining	.	.	1.07	1.54	1.10	1.38
All Median Combining	.	.	0.94	0.83	1.01	1.04
All Mean Combining	.	.	0.91	0.85	1.00	0.98
All Tr. Mean Combining	.	.	0.92	0.82	1.00	1.00
All MSE(A) Combining	.	.	.	0.86	1.03	1.02
All MSE(B) Combining	.	.	.	0.84	1.03	1.01
All MSE(C) Combining	.	.	.	0.84	1.02	1.01
All MSE(D) Combining	.	.	.	0.88	1.05	1.04
All MSE(E) Combining	.	.	.	0.85	1.05	1.03
All MSE(F) Combining	.	.	.	0.84	1.04	1.02
All Rec. Best(4q) Combining	.	.	1.19	1.53	1.17	1.08
All Rec. Best(8q) Combining	.	.	0.96	1.67	1.45	1.50
UCSV and Triangle Rec. Best(4q) Combining	.	.	.	1.37	1.06	1.00
UCSV and Triangle Rec. Best(8q) Combining	.	.	.	1.02	1.16	1.09

Source: Authors' calculations.

Table 3.3
Root Mean Squared Errors for Inflation Forecasting Models by Subperiod, Relative to the Unobserved Components-Stochastic Volatility Model: PCE-all

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
Number of observations	32	36	32	32	32	25
Root MSE of UC-SV forecast	0.73	1.83	1.41	0.88	0.59	0.72
Forecasting model and relative root mean square errors						
Univariate forecasts						
UC-SV	1.00	1.00	1.00	1.00	1.00	1.00
AR(AIC)_rec	.	1.14	1.02	1.15	1.06	1.45
AR(AIC)_iter_rec	.	1.04	1.01	1.14	1.04	1.50
AR(BIC)_rec	.	1.12	1.02	1.15	1.07	1.58
AO	1.02	1.20	1.18	1.01	1.10	1.09
MA(1)_rec	.	1.08	1.00	1.08	1.04	1.42
AR(4)_rec	.	1.16	1.03	1.13	1.07	1.46
AR(AIC)_roll	.	1.13	1.06	1.09	1.25	1.28
AR(AIC)_iter_roll	.	1.26	1.06	1.20	1.31	1.35
AR(BIC)_roll	.	1.11	1.07	1.09	1.24	1.33
AR(4)_roll	.	1.23	1.06	1.15	1.22	1.26
AR(24)_iter	.	.	1.23	1.53	1.15	1.34
AR(24)_iter_nocon	.	.	1.10	1.42	1.12	1.33
MA(1)_roll	.	1.04	0.99	1.09	1.03	1.10
MA(2) - NS	1.01	1.09	1.20	1.00	1.01	1.15
MA(1), θ=.25	1.10	1.01	0.99	1.12	1.07	1.59
MA(1), θ=.65	0.99	1.12	1.18	0.99	1.02	1.14
Single-predictor ADL forecasts						
UR(Level)_AIC_rec	.	1.06	0.98	0.99	1.22	1.43
UR(Dif)_AIC_rec	.	1.02	1.04	1.10	1.15	1.47
UR(1sdBP)_AIC_rec	.	1.07	1.09	0.99	1.14	1.42
GDP(Dif)_AIC_rec	.	1.02	0.98	1.04	1.15	1.47
GDP(1sdBP)_AIC_rec	.	1.08	1.02	1.00	1.11	1.43
IP(Dif)_AIC_rec	.	1.01	1.00	1.04	1.24	1.51
IP(1sdBP)_AIC_rec	.	1.06	1.06	1.01	1.21	1.46
Emp(Dif)_AIC_rec	.	1.03	0.96	1.04	1.12	1.66
Emp(1sdBP)_AIC_rec	.	1.04	1.03	1.01	1.13	1.54
CapU(Level)_AIC_rec	.	.	.	1.13	1.31	1.75
CapU((Dif)_AIC_rec	.	.	.	1.21	1.31	1.70
CapU(1sdBP)_AIC_rec	.	.	.	1.13	1.22	1.50
HPerm(Level)_AIC_rec	.	.	0.96	1.05	1.07	1.74
HPerm((Dif)_AIC_rec	.	.	1.11	1.16	1.07	1.61
HPerm(1sdBP)_AIC_rec	.	.	0.99	1.03	1.07	1.43
CFNAI(Dif)_AIC_rec	.	.	.	1.15	1.18	1.76
CFNAI(1sdBP)_AIC_rec	.	.	.	1.12	1.19	1.64
UR_5wk(Level)_AIC_rec	.	1.19	0.95	1.13	1.47	1.44
UR_5wk(Dif)_AIC_rec	.	1.07	1.02	1.14	1.19	1.45
UR_5wk(1sdBP)_AIC_rec	.	1.09	0.98	1.12	1.19	1.38
AHE(Dif)_AIC_rec	.	.	1.10	1.27	1.08	1.59

Table 3.3 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
AHE(1sdBP)_AIC_rec	.	.	1.15	1.26	1.07	1.56
RealAHE(Dif)_AIC_rec	.	.	1.10	1.27	1.08	1.59
RealAHE(1sdBP)_AIC_rec	.	.	1.15	1.26	1.07	1.56
LaborShare(Level)_AIC_rec	.	1.15	0.96	1.28	1.65	1.45
LaborShare(Dif)_AIC_rec	.	1.15	1.01	1.14	1.09	1.40
ULaborShare(1sdBP)_AIC_rec	.	1.18	0.96	1.13	1.32	1.40
CPI_Med(Level)_AIC_rec	.	.	.	1.33	1.34	1.56
CPI_Med(Dif)_AIC_rec	.	.	.	1.22	1.13	1.64
CPI_TrMn(Level)_AIC_rec	.	.	.	1.28	1.34	1.50
CPI_TrMn(Dif)_AIC_rec	.	.	.	1.15	1.10	1.64
ExRate(Dif)_AIC_rec	.	.	.	1.28	1.33	1.31
ExRate(1sdBP)_AIC_rec	.	.	.	1.60	1.15	1.32
tb_spr_AIC_rec	.	1.18	1.31	1.25	1.01	1.51
UR(Level)_AIC_roll	.	1.24	1.53	1.04	1.32	1.25
UR(Dif)_AIC_roll	.	1.10	1.12	1.11	1.26	1.21
UR(1sdBP)_AIC_roll	.	1.29	1.30	1.06	1.29	1.25
GDP(Dif)_AIC_roll	.	1.07	1.14	1.16	1.33	1.19
GDP(1sdBP)_AIC_roll	.	1.24	1.12	1.02	1.33	1.21
IP(Dif)_AIC_roll	.	1.08	0.99	1.19	1.39	1.25
IP(1sdBP)_AIC_roll	.	1.20	1.18	1.14	1.43	1.37
Emp(Dif)_AIC_roll	.	1.10	1.20	1.12	1.34	1.22
Emp(1sdBP)_AIC_roll	.	1.24	1.12	1.15	1.34	1.32
CapU(Level)_AIC_roll	.	.	.	1.10	1.44	1.29
CapU((Dif)_AIC_roll	.	.	.	1.17	1.37	1.30
CapU(1sdBP)_AIC_roll	.	.	.	1.01	1.43	1.30
HPerm(Level)_AIC_roll	.	.	1.00	1.14	1.25	1.33
HPerm((Dif)_AIC_roll	.	.	1.19	1.08	1.11	1.88
HPerm(1sdBP)_AIC_roll	.	.	1.07	1.15	1.28	1.24
CFNAI(Dif)_AIC_roll	.	.	.	1.14	1.38	1.21
CFNAI(1sdBP)_AIC_roll	.	.	.	1.10	1.37	1.28
UR_5wk(Level)_AIC_roll	.	1.32	1.25	1.16	1.43	1.26
UR_5wk(Dif)_AIC_roll	.	1.12	1.06	1.17	1.34	1.28
UR_5wk(1sdBP)_AIC_roll	.	1.20	1.10	1.03	1.43	1.27
AHE(Dif)_AIC_roll	.	.	1.10	1.35	1.33	1.24
AHE(1sdBP)_AIC_roll	.	.	1.15	1.06	1.34	1.14
RealAHE(Dif)_AIC_roll	.	.	1.10	1.35	1.33	1.24
RealAHE(1sdBP)_AIC_roll	.	.	1.15	1.06	1.34	1.14
LaborShare(Level)_AIC_roll	.	1.26	1.01	1.21	1.45	1.27
LaborShare(Dif)_AIC_roll	.	1.16	1.04	1.11	1.32	1.48
ULaborShare(1sdBP)_AIC_roll	.	1.22	1.03	1.09	1.35	1.23
CPI_Med(Level)_AIC_roll	.	.	.	1.16	1.27	1.20
CPI_Med(Dif)_AIC_roll	.	.	.	1.04	1.25	1.32
CPI_TrMn(Level)_AIC_roll	.	.	.	1.23	1.31	1.13
CPI_TrMn(Dif)_AIC_roll	.	.	.	1.07	1.29	1.28
ExRate(Dif)_AIC_roll	.	.	.	1.33	1.25	1.24
ExRate(1sdBP)_AIC_roll	.	.	.	1.64	1.25	1.22

Table 3.3 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
tb_spr_AIC_roll	.	1.18	1.79	1.39	1.38	1.35
UR(Level)_BIC_rec	.	1.05	0.98	1.06	1.22	1.42
UR(Dif)_BIC_rec	.	1.00	1.08	1.11	1.15	1.48
UR(1sdBP)_BIC_rec	.	1.05	1.07	1.06	1.14	1.42
GDP(Dif)_BIC_rec	.	1.03	1.09	1.10	1.15	1.47
GDP(1sdBP)_BIC_rec	.	1.09	0.99	1.07	1.07	1.44
IP(Dif)_BIC_rec	.	1.04	1.03	1.10	1.23	1.50
IP(1sdBP)_BIC_rec	.	1.07	1.02	1.09	1.16	1.47
Emp(Dif)_BIC_rec	.	1.02	0.97	1.02	1.11	1.59
Emp(1sdBP)_BIC_rec	.	1.06	0.98	1.06	1.13	1.54
CapU(Level)_BIC_rec	.	.	.	1.16	1.28	1.75
CapU((Dif)_BIC_rec	.	.	.	1.22	1.31	1.79
CapU(1sdBP)_BIC_rec	.	.	.	1.10	1.20	1.47
HPerm(Level)_BIC_rec	.	.	0.95	1.12	1.07	1.74
HPerm((Dif)_BIC_rec	.	.	1.07	1.21	1.07	1.61
HPerm(1sdBP)_BIC_rec	.	.	0.97	1.09	1.07	1.43
CFNAI(Dif)_BIC_rec	.	.	.	1.16	1.17	1.81
CFNAI(1sdBP)_BIC_rec	.	.	.	1.09	1.18	1.54
UR_5wk(Level)_BIC_rec	.	1.16	0.94	1.15	1.43	1.57
UR_5wk(Dif)_BIC_rec	.	1.02	1.02	1.15	1.12	1.54
UR_5wk(1sdBP)_BIC_rec	.	1.08	0.97	1.13	1.20	1.43
AHE(Dif)_BIC_rec	.	.	1.07	1.29	1.11	1.62
AHE(1sdBP)_BIC_rec	.	.	1.15	1.31	1.08	1.57
RealAHE(Dif)_BIC_rec	.	.	1.07	1.29	1.11	1.62
RealAHE(1sdBP)_BIC_rec	.	.	1.15	1.31	1.08	1.57
LaborShare(Level)_BIC_rec	.	1.15	0.98	1.30	1.51	1.65
LaborShare(Dif)_BIC_rec	.	1.11	1.01	1.14	1.10	1.54
ULaborShare(1sdBP)_BIC_rec	.	1.18	0.99	1.14	1.32	1.64
CPI_Med(Level)_BIC_rec	.	.	.	1.37	1.26	1.63
CPI_Med(Dif)_BIC_rec	.	.	.	1.27	1.13	1.64
CPI_TrMn(Level)_BIC_rec	.	.	.	1.35	1.25	1.61
CPI_TrMn(Dif)_BIC_rec	.	.	.	1.27	1.10	1.64
ExRate(Dif)_BIC_rec	.	.	.	1.38	1.25	1.41
ExRate(1sdBP)_BIC_rec	.	.	.	1.53	1.16	1.46
tb_spr_BIC_rec	.	1.14	1.13	1.16	1.04	1.59
UR(Level)_BIC_roll	.	1.23	1.29	1.10	1.32	1.31
UR(Dif)_BIC_roll	.	1.06	1.13	1.08	1.33	1.30
UR(1sdBP)_BIC_roll	.	1.22	1.29	1.10	1.32	1.31
GDP(Dif)_BIC_roll	.	1.05	1.12	1.09	1.33	1.32
GDP(1sdBP)_BIC_roll	.	1.23	1.09	1.12	1.28	1.30
IP(Dif)_BIC_roll	.	1.05	1.01	1.18	1.43	1.27
IP(1sdBP)_BIC_roll	.	1.23	1.17	1.15	1.46	1.39
Emp(Dif)_BIC_roll	.	1.09	1.12	1.14	1.36	1.30
Emp(1sdBP)_BIC_roll	.	1.18	1.22	1.18	1.39	1.29
CapU(Level)_BIC_roll	.	.	.	1.15	1.48	1.34
CapU(Dif)_BIC_roll	.	.	.	1.16	1.40	1.32

Table 3.3 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
CapU(1sdBP)_BIC_roll	.	.	.	1.09	1.46	1.36
HPerm(Level)_BIC_roll	.	.	1.00	1.10	1.27	1.39
HPerm((Dif)_BIC_roll	.	.	1.09	1.07	1.10	1.33
HPerm(1sdBP)_BIC_roll	.	.	1.00	1.09	1.27	1.33
CFNAI(Dif)_BIC_roll	.	.	.	1.13	1.41	1.27
CFNAI(1sdBP)_BIC_roll	.	.	.	1.06	1.39	1.28
UR_5wk(Level)_BIC_roll	.	1.26	1.24	1.19	1.42	1.33
UR_5wk(Dif)_BIC_roll	.	1.04	1.07	1.12	1.32	1.30
UR_5wk(1sdBP)_BIC_roll	.	1.17	1.10	1.06	1.42	1.27
AHE(Dif)_BIC_roll	.	.	1.08	1.09	1.27	1.30
AHE(1sdBP)_BIC_roll	.	.	1.15	1.13	1.33	1.26
RealAHE(Dif)_BIC_roll	.	.	1.08	1.09	1.27	1.30
RealAHE(1sdBP)_BIC_roll	.	.	1.15	1.13	1.33	1.26
LaborShare(Level)_BIC_roll	.	1.19	0.97	1.20	1.38	1.31
LaborShare(Dif)_BIC_roll	.	1.13	1.04	1.24	1.34	1.44
ULaborShare(1sdBP)_BIC_roll	.	1.17	0.96	1.11	1.35	1.32
CPI_Med(Level)_BIC_roll	.	.	.	1.12	1.29	1.19
CPI_Med(Dif)_BIC_roll	.	.	.	1.09	1.27	1.29
CPI_TrMn(Level)_BIC_roll	.	.	.	1.15	1.27	1.16
CPI_TrMn(Dif)_BIC_roll	.	.	.	1.08	1.28	1.34
ExRate(Dif)_BIC_roll	.	.	.	1.28	1.25	1.40
ExRate(1sdBP)_BIC_roll	.	.	.	1.48	1.26	1.40
tb_spr_BIC_roll	.	1.15	1.76	1.15	1.33	1.37
Triangle model forecasts						
Triangle Constant NAIRU	.	.	1.14	1.18	1.25	1.20
Triangle TV NAIRU	.	.	1.07	1.20	1.04	1.30
Triangle Constant NAIRU (no z)	.	.	0.98	1.33	1.38	1.27
Triangle TV NAIRU (no z)	.	.	0.97	1.48	1.16	1.58
Combination forecasts						
Activity Median Combining	.	1.07	0.97	1.05	1.16	1.32
Activity Mean Combining	.	1.07	0.94	1.05	1.14	1.35
Activity Tr. Mean Combining	.	1.07	0.95	1.05	1.16	1.34
Activity MSE(A) Combining	.	.	0.95	1.04	1.15	1.35
Activity MSE(B) Combining	.	.	0.95	1.04	1.15	1.35
Activity MSE(C) Combining	.	.	0.95	1.04	1.14	1.35
Activity MSE(D) Combining	.	.	0.95	1.05	1.16	1.35
Activity MSE(E) Combining	.	.	0.95	1.04	1.16	1.35
Activity MSE(F) Combining	.	.	0.94	1.04	1.15	1.35
Activity Rec. Best(4q) Combining	.	1.20	1.10	1.25	1.44	1.42
Activity Rec. Best(8q) Combining	.	1.21	0.99	1.19	1.46	1.61
OtherADL Median Combining	.	1.14	1.04	1.15	1.17	1.33
OtherADL Mean Combining	.	1.14	1.02	1.12	1.14	1.35
OtherADL Tr. Mean Combining	.	1.14	1.03	1.13	1.14	1.35
OtherADL MSE(A) Combining	.	.	0.93	1.14	1.15	1.34
OtherADL MSE(B) Combining	.	.	0.93	1.14	1.15	1.34
OtherADL MSE(C) Combining	.	.	0.94	1.14	1.16	1.35

Table 3.3 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
OtherADL MSE(D) Combining	.	.	0.94	1.14	1.16	1.32
OtherADL MSE(E) Combining	.	.	0.93	1.13	1.17	1.34
OtherADL MSE(F) Combining	.	.	0.94	1.13	1.17	1.36
OtherADL Rec. Best(4q) Combining	.	1.29	1.13	1.18	1.32	1.11
OtherADL Rec. Best(8q) Combining	.	1.32	0.98	1.32	1.33	1.26
All Median Combining	.	1.07	0.94	1.07	1.15	1.31
All Mean Combining	.	1.08	0.93	1.06	1.12	1.34
All Tr. Mean Combining	.	1.08	0.93	1.07	1.14	1.33
All MSE(A) Combining	.	.	0.92	1.06	1.14	1.34
All MSE(B) Combining	.	.	0.92	1.06	1.14	1.34
All MSE(C) Combining	.	.	0.92	1.06	1.13	1.34
All MSE(D) Combining	.	.	0.93	1.07	1.16	1.33
All MSE(E) Combining	.	.	0.92	1.06	1.15	1.34
All MSE(F) Combining	.	.	0.92	1.06	1.14	1.34
All Rec. Best(4q) Combining	.	1.25	1.18	1.34	1.44	1.32
All Rec. Best(8q) Combining	.	1.22	1.01	1.30	1.52	1.44
UCSV and Triangle Rec. Best(4q) Combining	.	.	.	1.07	1.18	1.07
UCSV and Triangle Rec. Best(8q) Combining	.	.	.	1.15	1.16	1.07

Source: Authors' calculations.

Table 3.4
Root Mean Squared Errors for Inflation Forecasting Models by Subperiod, Relative to the Unobserved Components-Stochastic Volatility Model: PCE-core

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
Number of observations	32	36	32	32	32	25
Root MSE of UC-SV forecast	0.68	1.56	1.08	0.55	0.36	0.33
Forecasting model and relative root mean square errors						
Univariate forecasts						
UC-SV	1.00	1.00	1.00	1.00	1.00	1.00
AR(AIC)_rec	.	.	1.15	1.15	1.21	1.34
AR(AIC)_iter_rec	.	.	1.09	1.17	1.24	1.37
AR(BIC)_rec	.	.	1.13	1.22	1.21	1.34
AO	1.08	1.16	1.12	1.00	0.94	1.18
MA(1)_rec	.	1.03	1.03	1.09	1.16	1.27
AR(4)_rec	.	.	1.15	1.15	1.18	1.29
AR(AIC)_roll	.	.	1.15	1.24	1.18	1.30
AR(AIC)_iter_roll	.	.	1.15	1.31	1.21	1.27
AR(BIC)_roll	.	.	1.15	1.06	1.24	1.28
AR(4)_roll	.	.	1.19	1.25	1.16	1.27
AR(24)_iter	.	.	.	1.85	1.37	1.26
AR(24)_iter_nocon	.	.	1.14	1.32	1.26	1.24
MA(1)_roll	.	1.02	1.03	1.08	1.14	1.06
MA(2) - NS	1.09	1.05	1.14	0.99	1.07	1.06
MA(1), θ=.25	1.01	1.01	0.99	1.11	1.19	1.33
MA(1), θ=.65	1.08	1.07	1.12	0.98	1.03	1.07
Single-predictor ADL forecasts						
UR(Level)_AIC_rec	.	.	1.08	1.01	1.48	1.51
UR(Dif)_AIC_rec	.	.	1.14	1.16	1.42	1.30
UR(1sdBP)_AIC_rec	.	.	1.19	1.08	1.30	1.41
GDP(Dif)_AIC_rec	.	.	1.03	1.17	1.38	1.45
GDP(1sdBP)_AIC_rec	.	.	0.95	1.01	1.06	1.50
IP(Dif)_AIC_rec	.	.	1.01	1.15	1.39	1.29
IP(1sdBP)_AIC_rec	.	.	0.96	1.01	1.25	1.80
Emp(Dif)_AIC_rec	.	.	1.00	1.14	1.26	2.10
Emp(1sdBP)_AIC_rec	.	.	1.14	1.06	1.28	1.70
CapU(Level)_AIC_rec	.	.	.	1.23	1.53	2.81
CapU((Dif)_AIC_rec	.	.	.	1.23	1.39	1.27
CapU(1sdBP)_AIC_rec	.	.	.	1.19	1.17	1.52
HPerm(Level)_AIC_rec	.	.	1.13	1.17	1.30	2.11
HPerm((Dif)_AIC_rec	.	.	1.25	1.17	1.22	1.34
HPerm(1sdBP)_AIC_rec	.	.	1.15	1.04	1.38	1.32
CFNAI(Dif)_AIC_rec	.	.	.	1.15	1.25	1.68
CFNAI(1sdBP)_AIC_rec	.	.	.	1.00	1.12	1.63
UR_5wk(Level)_AIC_rec	.	.	0.96	1.18	1.99	1.73
UR_5wk(Dif)_AIC_rec	.	.	1.14	1.21	1.44	1.21
UR_5wk(1sdBP)_AIC_rec	.	.	1.06	1.22	1.57	1.24
AHE(Dif)_AIC_rec	.	.	1.18	1.24	1.22	1.34

Table 3.4 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
AHE(1sdBP)_AIC_rec	.	.	1.39	1.29	1.27	1.34
RealAHE(Dif)_AIC_rec	.	.	1.18	1.24	1.22	1.34
RealAHE(1sdBP)_AIC_rec	.	.	1.39	1.29	1.27	1.34
LaborShare(Level)_AIC_rec	.	.	0.99	1.39	1.91	1.57
LaborShare(Dif)_AIC_rec	.	.	1.09	1.18	1.40	1.50
ULaborShare(1sdBP)_AIC_rec	.	.	1.05	1.25	1.55	1.59
CPI_Med(Level)_AIC_rec	.	.	.	1.20	1.33	1.37
CPI_Med(Dif)_AIC_rec	.	.	.	1.12	1.17	1.49
CPI_TrMn(Level)_AIC_rec	.	.	.	1.20	1.19	1.32
CPI_TrMn(Dif)_AIC_rec	.	.	.	1.09	1.05	1.28
ExRate(Dif)_AIC_rec	.	.	.	1.29	1.35	1.27
ExRate(1sdBP)_AIC_rec	.	.	.	1.45	1.26	1.28
tb_spr_AIC_rec	.	.	1.40	1.33	1.22	1.49
UR(Level)_AIC_roll	.	.	1.44	1.20	1.03	1.49
UR(Dif)_AIC_roll	.	.	1.22	1.20	1.20	1.31
UR(1sdBP)_AIC_roll	.	.	1.20	0.93	1.01	1.60
GDP(Dif)_AIC_roll	.	.	1.17	1.24	1.23	1.33
GDP(1sdBP)_AIC_roll	.	.	1.00	0.98	1.01	1.55
IP(Dif)_AIC_roll	.	.	1.10	1.35	1.29	1.44
IP(1sdBP)_AIC_roll	.	.	0.98	1.11	1.20	1.73
Emp(Dif)_AIC_roll	.	.	1.08	1.14	1.18	1.58
Emp(1sdBP)_AIC_roll	.	.	1.23	1.17	1.12	2.07
CapU(Level)_AIC_roll	.	.	.	1.21	1.33	1.73
CapU((Dif)_AIC_roll	.	.	.	1.34	1.27	1.44
CapU(1sdBP)_AIC_roll	.	.	.	1.00	1.26	1.60
HPerm(Level)_AIC_roll	.	.	1.27	1.17	1.03	1.51
HPerm((Dif)_AIC_roll	.	.	1.37	1.17	1.26	1.23
HPerm(1sdBP)_AIC_roll	.	.	1.44	1.11	1.15	1.32
CFNAI(Dif)_AIC_roll	.	.	.	1.14	1.20	1.49
CFNAI(1sdBP)_AIC_roll	.	.	.	0.96	1.09	1.73
UR_5wk(Level)_AIC_roll	.	.	1.14	1.44	1.04	1.44
UR_5wk(Dif)_AIC_roll	.	.	1.17	1.06	1.16	1.28
UR_5wk(1sdBP)_AIC_roll	.	.	1.13	1.00	1.26	1.33
AHE(Dif)_AIC_roll	.	.	1.18	1.25	1.21	1.32
AHE(1sdBP)_AIC_roll	.	.	1.36	1.36	1.21	1.61
RealAHE(Dif)_AIC_roll	.	.	1.18	1.25	1.21	1.32
RealAHE(1sdBP)_AIC_roll	.	.	1.36	1.36	1.21	1.61
LaborShare(Level)_AIC_roll	.	.	1.20	1.35	1.56	1.58
LaborShare(Dif)_AIC_roll	.	.	1.32	1.12	1.16	1.63
ULaborShare(1sdBP)_AIC_roll	.	.	1.27	1.23	1.46	1.55
CPI_Med(Level)_AIC_roll	.	.	.	1.37	1.31	1.14
CPI_Med(Dif)_AIC_roll	.	.	.	1.20	1.21	1.31
CPI_TrMn(Level)_AIC_roll	.	.	.	1.34	1.28	1.33
CPI_TrMn(Dif)_AIC_roll	.	.	.	1.15	1.19	1.33
ExRate(Dif)_AIC_roll	.	.	.	1.39	1.21	1.44
ExRate(1sdBP)_AIC_roll	.	.	.	1.43	1.12	1.46

Table 3.4 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
tb_spr_AIC_roll	.	.	1.64	1.26	1.24	1.30
UR(Level)_BIC_rec	.	.	0.96	1.06	1.48	1.58
UR(Dif)_BIC_rec	.	.	1.17	1.16	1.42	1.34
UR(1sdBP)_BIC_rec	.	.	1.10	1.12	1.35	1.39
GDP(Dif)_BIC_rec	.	.	1.10	1.17	1.38	1.47
GDP(1sdBP)_BIC_rec	.	.	0.95	1.07	1.14	1.46
IP(Dif)_BIC_rec	.	.	1.14	1.15	1.40	1.33
IP(1sdBP)_BIC_rec	.	.	0.97	1.03	1.23	1.67
Emp(Dif)_BIC_rec	.	.	1.01	1.15	1.30	2.09
Emp(1sdBP)_BIC_rec	.	.	1.06	1.10	1.27	1.66
CapU(Level)_BIC_rec	.	.	.	1.25	1.45	2.53
CapU((Dif)_BIC_rec	.	.	.	1.23	1.39	1.27
CapU(1sdBP)_BIC_rec	.	.	.	1.20	1.18	1.48
HPerm(Level)_BIC_rec	.	.	1.10	1.13	1.20	1.74
HPerm((Dif)_BIC_rec	.	.	1.15	1.23	1.22	1.34
HPerm(1sdBP)_BIC_rec	.	.	1.11	1.11	1.38	1.38
CFNAI(Dif)_BIC_rec	.	.	.	1.16	1.30	1.68
CFNAI(1sdBP)_BIC_rec	.	.	.	1.09	1.16	1.63
UR_5wk(Level)_BIC_rec	.	.	1.03	1.18	1.98	1.76
UR_5wk(Dif)_BIC_rec	.	.	1.14	1.20	1.21	1.25
UR_5wk(1sdBP)_BIC_rec	.	.	1.04	1.23	1.58	1.31
AHE(Dif)_BIC_rec	.	.	1.19	1.26	1.22	1.34
AHE(1sdBP)_BIC_rec	.	.	1.31	1.37	1.31	1.28
RealAHE(Dif)_BIC_rec	.	.	1.19	1.26	1.22	1.34
RealAHE(1sdBP)_BIC_rec	.	.	1.31	1.37	1.31	1.28
LaborShare(Level)_BIC_rec	.	.	1.07	1.32	1.73	1.40
LaborShare(Dif)_BIC_rec	.	.	1.06	1.14	1.46	1.36
ULaborShare(1sdBP)_BIC_rec	.	.	1.08	1.22	1.66	1.46
CPI_Med(Level)_BIC_rec	.	.	.	1.29	1.36	1.46
CPI_Med(Dif)_BIC_rec	.	.	.	1.19	1.24	1.48
CPI_TrMn(Level)_BIC_rec	.	.	.	1.21	1.28	1.38
CPI_TrMn(Dif)_BIC_rec	.	.	.	1.14	1.19	1.33
ExRate(Dif)_BIC_rec	.	.	.	1.15	1.38	1.31
ExRate(1sdBP)_BIC_rec	.	.	.	1.19	1.26	1.32
tb_spr_BIC_rec	.	.	1.26	1.22	1.12	1.50
UR(Level)_BIC_roll	.	.	1.40	1.18	1.01	1.50
UR(Dif)_BIC_roll	.	.	1.25	1.02	1.21	1.30
UR(1sdBP)_BIC_roll	.	.	1.27	0.97	1.00	1.57
GDP(Dif)_BIC_roll	.	.	1.17	1.10	1.23	1.30
GDP(1sdBP)_BIC_roll	.	.	1.02	1.02	1.00	1.53
IP(Dif)_BIC_roll	.	.	1.24	1.11	1.31	1.34
IP(1sdBP)_BIC_roll	.	.	1.05	1.12	1.18	1.65
Emp(Dif)_BIC_roll	.	.	1.14	1.11	1.24	1.47
Emp(1sdBP)_BIC_roll	.	.	1.17	1.06	1.12	2.04
CapU(Level)_BIC_roll	.	.	.	1.21	1.33	1.44
CapU(Dif)_BIC_roll	.	.	.	1.07	1.28	1.39

Table 3.4 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
CapU(1sdBP)_BIC_roll	.	.	.	1.07	1.24	1.37
HPerm(Level)_BIC_roll	.	.	1.15	1.15	1.00	1.54
HPerm((Dif)_BIC_roll	.	.	1.19	1.08	1.26	1.21
HPerm(1sdBP)_BIC_roll	.	.	1.20	1.12	1.11	1.31
CFNAI(Dif)_BIC_roll	.	.	.	1.08	1.24	1.37
CFNAI(1sdBP)_BIC_roll	.	.	.	0.99	1.09	1.74
UR_5wk(Level)_BIC_roll	.	.	1.14	1.42	1.05	1.45
UR_5wk(Dif)_BIC_roll	.	.	1.20	1.05	1.18	1.28
UR_5wk(1sdBP)_BIC_roll	.	.	1.16	1.01	1.27	1.27
AHE(Dif)_BIC_roll	.	.	1.16	1.11	1.24	1.34
AHE(1sdBP)_BIC_roll	.	.	1.31	1.17	1.22	1.49
RealAHE(Dif)_BIC_roll	.	.	1.16	1.11	1.24	1.34
RealAHE(1sdBP)_BIC_roll	.	.	1.31	1.17	1.22	1.49
LaborShare(Level)_BIC_roll	.	.	1.14	1.22	1.41	1.45
LaborShare(Dif)_BIC_roll	.	.	1.12	1.09	1.26	1.58
ULaborShare(1sdBP)_BIC_roll	.	.	1.24	1.15	1.45	1.46
CPI_Med(Level)_BIC_roll	.	.	.	1.24	1.29	1.14
CPI_Med(Dif)_BIC_roll	.	.	.	1.05	1.22	1.27
CPI_TrMn(Level)_BIC_roll	.	.	.	1.23	1.29	1.30
CPI_TrMn(Dif)_BIC_roll	.	.	.	1.07	1.25	1.31
ExRate(Dif)_BIC_roll	.	.	.	1.14	1.29	1.43
ExRate(1sdBP)_BIC_roll	.	.	.	1.07	1.08	1.41
tb_spr_BIC_roll	.	.	1.62	1.10	1.25	1.27
Triangle model forecasts						
Triangle Constant NAIRU	.	.	1.69	1.06	1.80	1.55
Triangle TV NAIRU	.	.	1.94	0.99	1.37	1.44
Triangle Constant NAIRU (no z)	.	.	1.17	1.64	2.20	1.58
Triangle TV NAIRU (no z)	.	.	1.28	1.58	1.48	2.13
Combination forecasts						
Activity Median Combining	.	.	0.95	1.01	1.07	1.27
Activity Mean Combining	.	.	0.93	1.00	1.07	1.28
Activity Tr. Mean Combining	.	.	0.95	1.01	1.08	1.27
Activity MSE(A) Combining	.	.	.	0.99	1.08	1.31
Activity MSE(B) Combining	.	.	.	0.98	1.08	1.30
Activity MSE(C) Combining	.	.	.	0.98	1.07	1.30
Activity MSE(D) Combining	.	.	.	0.97	1.10	1.34
Activity MSE(E) Combining	.	.	.	0.95	1.08	1.33
Activity MSE(F) Combining	.	.	.	0.95	1.07	1.33
Activity Rec. Best(4q) Combining	.	.	1.01	1.39	1.37	1.95
Activity Rec. Best(8q) Combining	.	.	0.78	1.35	1.29	1.57
OtherADL Median Combining	.	.	1.10	1.11	1.16	1.24
OtherADL Mean Combining	.	.	1.08	1.12	1.14	1.23
OtherADL Tr. Mean Combining	.	.	1.10	1.11	1.14	1.22
OtherADL MSE(A) Combining	.	.	.	1.13	1.16	1.26
OtherADL MSE(B) Combining	.	.	.	1.13	1.16	1.25
OtherADL MSE(C) Combining	.	.	.	1.13	1.16	1.24

Table 3.4 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
OtherADL MSE(D) Combining	.	.	.	1.13	1.18	1.29
OtherADL MSE(E) Combining	.	.	.	1.13	1.17	1.26
OtherADL MSE(F) Combining	.	.	.	1.13	1.17	1.25
OtherADL Rec. Best(4q) Combining	.	.	1.30	1.37	1.51	1.42
OtherADL Rec. Best(8q) Combining	.	.	1.40	1.32	1.49	1.72
All Median Combining	.	.	1.01	1.03	1.09	1.26
All Mean Combining	.	.	0.96	1.03	1.08	1.24
All Tr. Mean Combining	.	.	0.98	1.04	1.09	1.24
All MSE(A) Combining	.	.	.	1.01	1.10	1.27
All MSE(B) Combining	.	.	.	1.01	1.09	1.26
All MSE(C) Combining	.	.	.	1.00	1.08	1.26
All MSE(D) Combining	.	.	.	1.00	1.12	1.30
All MSE(E) Combining	.	.	.	0.98	1.10	1.29
All MSE(F) Combining	.	.	.	0.97	1.08	1.28
All Rec. Best(4q) Combining	.	.	1.20	1.44	1.43	1.93
All Rec. Best(8q) Combining	.	.	0.78	1.35	1.43	1.61
UCSV and Triangle Rec. Best(4q) Combining	.	.	.	1.04	1.13	1.03
UCSV and Triangle Rec. Best(8q) Combining	.	.	.	1.04	1.26	1.13

Source: Authors' calculations.

Table 3.5
Root Mean Squared Errors for Inflation Forecasting Models by Subperiod, Relative to the Unobserved Components-Stochastic Volatility Model: GDP deflator

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
Number of observations	32	36	32	32	32	25
Root MSE of UC-SV forecast	0.72	1.76	1.28	0.70	0.41	0.57
Forecasting model and relative root mean square errors						
Univariate forecasts						
UC-SV	1.00	1.00	1.00	1.00	1.00	1.00
AR(AIC)_rec	.	1.03	1.06	1.06	1.02	1.16
AR(AIC)_iter_rec	.	1.11	1.08	1.04	0.99	1.19
AR(BIC)_rec	.	1.03	1.04	1.08	1.07	1.24
AO	0.97	1.10	1.17	1.04	0.95	1.02
MA(1)_rec	.	1.02	1.00	1.04	1.02	1.16
AR(4)_rec	.	1.07	1.07	1.04	0.99	1.17
AR(AIC)_roll	.	1.11	1.05	1.15	1.19	1.16
AR(AIC)_iter_roll	.	1.10	1.06	1.02	1.12	1.11
AR(BIC)_roll	.	1.08	1.05	1.21	1.17	1.11
AR(4)_roll	.	1.15	1.08	1.08	1.14	1.15
AR(24)_iter	.	.	1.42	1.10	1.02	0.99
AR(24)_iter_nocon	.	.	1.34	1.02	0.99	0.99
MA(1)_roll	.	1.03	0.99	1.05	0.98	1.02
MA(2) - NS	0.97	1.02	1.19	1.03	1.02	1.02
MA(1), θ=.25	1.03	1.00	1.00	1.07	1.08	1.25
MA(1), θ=.65	0.96	1.03	1.17	1.02	0.99	1.01
Single-predictor ADL forecasts						
UR(Level)_AIC_rec	.	0.93	0.99	0.91	1.23	1.30
UR(Dif)_AIC_rec	.	0.93	1.11	0.96	1.25	1.22
UR(1sdBP)_AIC_rec	.	0.94	1.12	0.91	1.14	1.19
GDP(Dif)_AIC_rec	.	0.94	1.04	0.91	1.06	1.09
GDP(1sdBP)_AIC_rec	.	0.98	1.01	0.89	0.96	1.15
IP(Dif)_AIC_rec	.	0.90	1.05	0.89	1.25	1.15
IP(1sdBP)_AIC_rec	.	0.93	1.04	0.86	1.18	1.23
Emp(Dif)_AIC_rec	.	0.93	1.03	0.93	1.11	1.42
Emp(1sdBP)_AIC_rec	.	0.94	1.05	0.94	1.19	1.35
CapU(Level)_AIC_rec	.	.	.	1.03	1.54	1.87
CapU((Dif)_AIC_rec	.	.	.	0.96	1.39	1.22
CapU(1sdBP)_AIC_rec	.	.	.	0.91	1.23	1.23
HPerm(Level)_AIC_rec	.	.	1.07	1.05	0.89	1.60
HPerm((Dif)_AIC_rec	.	.	1.17	1.09	1.01	1.13
HPerm(1sdBP)_AIC_rec	.	.	1.18	1.04	1.10	1.09
CFNAI(Dif)_AIC_rec	.	.	.	1.00	1.16	1.54
CFNAI(1sdBP)_AIC_rec	.	.	.	0.89	1.12	1.46
UR_5wk(Level)_AIC_rec	.	1.08	1.01	0.92	1.83	1.32
UR_5wk(Dif)_AIC_rec	.	0.99	1.07	0.93	1.32	1.11
UR_5wk(1sdBP)_AIC_rec	.	0.99	1.04	0.95	1.30	1.14
AHE(Dif)_AIC_rec	.	.	1.09	1.09	1.05	1.16

Table 3.5 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
AHE(1sdBP)_AIC_rec	.	.	1.21	1.19	1.07	1.13
RealAHE(Dif)_AIC_rec	.	.	1.09	1.09	1.05	1.16
RealAHE(1sdBP)_AIC_rec	.	.	1.21	1.19	1.07	1.13
LaborShare(Level)_AIC_rec	.	1.10	0.99	1.18	1.85	1.27
LaborShare(Dif)_AIC_rec	.	1.09	1.06	1.07	1.12	1.14
ULaborShare(1sdBP)_AIC_rec	.	1.10	1.01	1.08	1.47	1.13
CPI_Med(Level)_AIC_rec	.	.	.	1.09	1.38	1.29
CPI_Med(Dif)_AIC_rec	.	.	.	1.02	1.15	1.34
CPI_TrMn(Level)_AIC_rec	.	.	.	1.12	1.17	1.21
CPI_TrMn(Dif)_AIC_rec	.	.	.	1.02	0.96	1.19
ExRate(Dif)_AIC_rec	.	.	.	1.23	1.44	1.06
ExRate(1sdBP)_AIC_rec	.	.	.	1.66	1.25	0.91
tb_spr_AIC_rec	.	1.08	1.23	1.19	0.95	1.27
UR(Level)_AIC_roll	.	1.09	1.43	1.06	1.16	1.11
UR(Dif)_AIC_roll	.	0.97	1.35	1.12	1.25	1.06
UR(1sdBP)_AIC_roll	.	1.03	1.33	1.04	1.25	1.13
GDP(Dif)_AIC_roll	.	1.13	1.21	1.08	1.26	1.05
GDP(1sdBP)_AIC_roll	.	1.15	0.96	1.05	1.21	1.07
IP(Dif)_AIC_roll	.	1.01	1.22	1.10	1.42	1.10
IP(1sdBP)_AIC_roll	.	1.05	1.10	1.12	1.48	1.20
Emp(Dif)_AIC_roll	.	1.04	1.24	1.15	1.23	1.08
Emp(1sdBP)_AIC_roll	.	1.06	1.23	1.17	1.27	1.23
CapU(Level)_AIC_roll	.	.	.	1.12	1.39	1.12
CapU((Dif)_AIC_roll	.	.	.	1.11	1.34	1.10
CapU(1sdBP)_AIC_roll	.	.	.	1.05	1.33	1.06
HPerm(Level)_AIC_roll	.	.	1.18	1.36	0.92	1.35
HPerm((Dif)_AIC_roll	.	.	1.20	1.21	1.12	1.08
HPerm(1sdBP)_AIC_roll	.	.	1.36	1.33	1.07	1.10
CFNAI(Dif)_AIC_roll	.	.	.	1.12	1.29	1.11
CFNAI(1sdBP)_AIC_roll	.	.	.	1.10	1.31	1.16
UR_5wk(Level)_AIC_roll	.	1.28	1.18	1.21	1.86	1.20
UR_5wk(Dif)_AIC_roll	.	1.09	1.11	1.17	1.45	1.14
UR_5wk(1sdBP)_AIC_roll	.	1.08	1.09	0.94	1.38	1.15
AHE(Dif)_AIC_roll	.	.	1.07	1.06	1.18	1.14
AHE(1sdBP)_AIC_roll	.	.	1.19	1.11	1.22	1.06
RealAHE(Dif)_AIC_roll	.	.	1.07	1.06	1.18	1.14
RealAHE(1sdBP)_AIC_roll	.	.	1.19	1.11	1.22	1.06
LaborShare(Level)_AIC_roll	.	1.22	1.39	1.18	1.25	1.14
LaborShare(Dif)_AIC_roll	.	1.11	1.53	1.11	1.60	1.26
ULaborShare(1sdBP)_AIC_roll	.	1.15	1.44	1.13	1.30	1.15
CPI_Med(Level)_AIC_roll	.	.	.	1.14	1.31	0.87
CPI_Med(Dif)_AIC_roll	.	.	.	1.06	1.20	1.14
CPI_TrMn(Level)_AIC_roll	.	.	.	1.16	1.21	0.72
CPI_TrMn(Dif)_AIC_roll	.	.	.	1.07	1.16	1.13
ExRate(Dif)_AIC_roll	.	.	.	1.42	1.15	1.19
ExRate(1sdBP)_AIC_roll	.	.	.	1.67	1.21	1.13

Table 3.5 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
tb_spr_AIC_roll	.	1.09	1.60	1.29	1.16	1.21
UR(Level)_BIC_rec	.	0.94	0.96	0.92	1.21	1.28
UR(Dif)_BIC_rec	.	0.94	1.08	0.97	1.20	1.20
UR(1sdBP)_BIC_rec	.	0.93	1.06	0.94	1.14	1.19
GDP(Dif)_BIC_rec	.	1.00	1.09	0.94	1.08	1.11
GDP(1sdBP)_BIC_rec	.	0.99	1.00	0.93	0.87	1.17
IP(Dif)_BIC_rec	.	0.96	1.06	0.89	1.22	1.11
IP(1sdBP)_BIC_rec	.	0.96	0.99	0.89	1.12	1.22
Emp(Dif)_BIC_rec	.	0.90	0.99	0.90	1.12	1.30
Emp(1sdBP)_BIC_rec	.	0.97	0.99	0.90	1.14	1.31
CapU(Level)_BIC_rec	.	.	.	1.06	1.47	1.83
CapU((Dif)_BIC_rec	.	.	.	1.02	1.39	1.22
CapU(1sdBP)_BIC_rec	.	.	.	0.94	1.19	1.21
HPerm(Level)_BIC_rec	.	.	0.99	1.05	0.89	1.60
HPerm((Dif)_BIC_rec	.	.	1.09	1.08	1.05	1.22
HPerm(1sdBP)_BIC_rec	.	.	1.08	0.99	1.09	1.05
CFNAI(Dif)_BIC_rec	.	.	.	1.00	1.16	1.51
CFNAI(1sdBP)_BIC_rec	.	.	.	0.86	1.08	1.32
UR_5wk(Level)_BIC_rec	.	1.08	1.00	1.01	1.73	1.43
UR_5wk(Dif)_BIC_rec	.	0.98	1.08	1.03	1.17	1.20
UR_5wk(1sdBP)_BIC_rec	.	1.00	1.03	0.98	1.28	1.18
AHE(Dif)_BIC_rec	.	.	1.07	1.17	1.07	1.24
AHE(1sdBP)_BIC_rec	.	.	1.21	1.31	1.11	1.20
RealAHE(Dif)_BIC_rec	.	.	1.07	1.17	1.07	1.24
RealAHE(1sdBP)_BIC_rec	.	.	1.21	1.31	1.11	1.20
LaborShare(Level)_BIC_rec	.	1.09	1.03	1.22	1.68	1.34
LaborShare(Dif)_BIC_rec	.	1.06	1.04	1.06	1.15	1.22
ULaborShare(1sdBP)_BIC_rec	.	1.11	1.03	1.09	1.49	1.23
CPI_Med(Level)_BIC_rec	.	.	.	1.14	1.38	1.29
CPI_Med(Dif)_BIC_rec	.	.	.	1.08	1.18	1.27
CPI_TrMn(Level)_BIC_rec	.	.	.	1.16	1.17	1.21
CPI_TrMn(Dif)_BIC_rec	.	.	.	1.08	0.99	1.22
ExRate(Dif)_BIC_rec	.	.	.	1.12	1.48	1.09
ExRate(1sdBP)_BIC_rec	.	.	.	1.28	1.33	1.11
tb_spr_BIC_rec	.	1.03	1.09	1.11	0.99	1.33
UR(Level)_BIC_roll	.	1.14	1.49	1.04	1.21	1.15
UR(Dif)_BIC_roll	.	1.03	1.33	1.14	1.21	1.10
UR(1sdBP)_BIC_roll	.	0.98	1.36	1.02	1.20	1.17
GDP(Dif)_BIC_roll	.	1.15	1.15	1.17	1.28	1.00
GDP(1sdBP)_BIC_roll	.	1.15	0.98	1.06	1.20	1.10
IP(Dif)_BIC_roll	.	1.05	1.16	1.10	1.34	1.08
IP(1sdBP)_BIC_roll	.	1.09	1.13	1.18	1.43	1.24
Emp(Dif)_BIC_roll	.	1.03	1.21	1.15	1.25	1.11
Emp(1sdBP)_BIC_roll	.	1.04	1.24	1.24	1.27	1.24
CapU(Level)_BIC_roll	.	.	.	1.01	1.37	1.14
CapU(Dif)_BIC_roll	.	.	.	1.10	1.27	1.05

Table 3.5 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
CapU(1sdBP)_BIC_roll	.	.	.	0.96	1.29	1.12
HPerm(Level)_BIC_roll	.	.	1.08	1.25	0.94	1.37
HPerm((Dif)_BIC_roll	.	.	1.08	1.25	1.14	1.07
HPerm(1sdBP)_BIC_roll	.	.	1.22	1.20	1.08	1.08
CFNAI(Dif)_BIC_roll	.	.	.	1.12	1.28	1.10
CFNAI(1sdBP)_BIC_roll	.	.	.	1.14	1.25	1.13
UR_5wk(Level)_BIC_roll	.	1.28	1.07	1.17	1.55	1.22
UR_5wk(Dif)_BIC_roll	.	1.07	1.08	1.14	1.27	1.08
UR_5wk(1sdBP)_BIC_roll	.	1.06	1.09	0.93	1.32	1.11
AHE(Dif)_BIC_roll	.	.	1.07	1.19	1.17	1.10
AHE(1sdBP)_BIC_roll	.	.	1.19	1.32	1.17	1.09
RealAHE(Dif)_BIC_roll	.	.	1.07	1.19	1.17	1.10
RealAHE(1sdBP)_BIC_roll	.	.	1.19	1.32	1.17	1.09
LaborShare(Level)_BIC_roll	.	1.21	1.24	1.22	1.24	1.06
LaborShare(Dif)_BIC_roll	.	1.06	1.32	1.23	1.31	1.16
ULaborShare(1sdBP)_BIC_roll	.	1.17	1.32	1.23	1.29	1.09
CPI_Med(Level)_BIC_roll	.	.	.	1.36	1.34	0.92
CPI_Med(Dif)_BIC_roll	.	.	.	1.13	1.18	1.26
CPI_TrMn(Level)_BIC_roll	.	.	.	1.26	1.14	0.73
CPI_TrMn(Dif)_BIC_roll	.	.	.	1.16	1.16	1.11
ExRate(Dif)_BIC_roll	.	.	.	1.26	1.22	1.16
ExRate(1sdBP)_BIC_roll	.	.	.	1.49	1.20	1.15
tb_spr_BIC_roll	.	1.13	1.50	1.21	1.14	1.17
Triangle model forecasts						
Triangle Constant NAIRU	.	.	1.08	0.78	1.22	1.20
Triangle TV NAIRU	.	.	0.98	0.81	1.07	1.23
Triangle Constant NAIRU (no z)	.	.	1.22	0.95	1.64	1.22
Triangle TV NAIRU (no z)	.	.	1.17	1.21	1.33	1.61
Combination forecasts						
Activity Median Combining	.	0.98	0.99	0.93	1.10	1.09
Activity Mean Combining	.	1.00	0.97	0.91	1.07	1.10
Activity Tr. Mean Combining	.	1.00	0.98	0.93	1.09	1.10
Activity MSE(A) Combining	.	.	0.98	0.91	1.09	1.11
Activity MSE(B) Combining	.	.	0.98	0.91	1.09	1.11
Activity MSE(C) Combining	.	.	0.98	0.91	1.08	1.10
Activity MSE(D) Combining	.	.	0.98	0.89	1.10	1.11
Activity MSE(E) Combining	.	.	0.98	0.89	1.09	1.11
Activity MSE(F) Combining	.	.	0.98	0.89	1.08	1.11
Activity Rec. Best(4q) Combining	.	1.09	1.24	1.02	1.33	1.07
Activity Rec. Best(8q) Combining	.	1.11	1.29	1.02	1.30	1.39
OtherADL Median Combining	.	1.09	1.02	1.06	1.12	1.08
OtherADL Mean Combining	.	1.10	0.99	1.02	1.09	1.07
OtherADL Tr. Mean Combining	.	1.09	0.98	1.05	1.09	1.07
OtherADL MSE(A) Combining	.	.	1.09	1.08	1.10	1.06
OtherADL MSE(B) Combining	.	.	1.07	1.08	1.11	1.06
OtherADL MSE(C) Combining	.	.	1.06	1.08	1.11	1.06

Table 3.5 (continued)

Forecast period	1960:Q1–1967:Q4	1968:Q1–1976:Q4	1977:Q1–1984:Q4	1985:Q1–1992:Q4	1993:Q1–2000:Q4	2001:Q1–2007:Q4
OtherADL MSE(D) Combining	.	.	1.11	1.08	1.10	1.04
OtherADL MSE(E) Combining	.	.	1.09	1.08	1.11	1.04
OtherADL MSE(F) Combining	.	.	1.07	1.08	1.12	1.06
OtherADL Rec. Best(4q) Combining	.	1.07	1.18	1.42	1.12	0.83
OtherADL Rec. Best(8q) Combining	.	1.15	1.32	1.22	1.05	0.91
All Median Combining	.	1.01	0.97	0.98	1.10	1.07
All Mean Combining	.	1.02	0.94	0.94	1.05	1.08
All Tr. Mean Combining	.	1.02	0.94	0.97	1.06	1.07
All MSE(A) Combining	.	.	0.97	0.95	1.07	1.07
All MSE(B) Combining	.	.	0.96	0.95	1.07	1.07
All MSE(C) Combining	.	.	0.96	0.95	1.07	1.07
All MSE(D) Combining	.	.	0.98	0.94	1.09	1.07
All MSE(E) Combining	.	.	0.98	0.93	1.09	1.07
All MSE(F) Combining	.	.	0.97	0.93	1.07	1.07
All Rec. Best(4q) Combining	.	1.07	1.35	1.20	1.32	0.90
All Rec. Best(8q) Combining	.	1.12	1.37	1.03	1.16	1.03
UCSV and Triangle Rec. Best(4q) Combining	.	.	.	0.91	1.14	1.17
UCSV and Triangle Rec. Best(8q) Combining	.	.	.	0.89	1.13	1.21

Source: Authors' calculations.

Notes

1. Experience has shown that the good in-sample fit of a forecasting model does not necessarily imply a good out-of-sample performance. The method of pseudo out-of-sample forecast evaluation aims to address this disjunction by simulating the experience a forecaster would have had using a forecasting model. In a pseudo out-of-sample forecasting exercise, one simulates standing at a given date t and performing all model specification and parameter estimation using only the data available at that date, then computing the h-period ahead forecast for date $t + h$; this is repeated for all dates in the forecast period.

2. A strict interpretation of pseudo out-of-sample forecasting would entail the use of real-time data (data of different vintages), but we interpret the term more generously to include the use of final data.

3. The specification in Gordon (1990), which is used here, differs from Gordon (1982, table 5, column 2) in three ways: (a) Gordon (1982) uses a polynomial distributed lag specification on lagged inflation, while Gordon (1990) uses a step function; (b) Gordon (1982) includes additional intercept shifts in 1970:Q3–1975:Q4 and 1976:Q1–1980:Q4, which are dropped in Gordon (1990); (c) Gordon (1982) uses Perry-weighted unemployment, whereas here we use overall unemployment.

4. Stockton and Glassman (1987), table 6, ratio of PHL(16,FE) to ARIMA RMSE for average of four intervals.

5. The random walk benchmark is a standard tool for forecast assessment, but it seems to have played at most a minor role in the inflation forecasting literature before Atkeson and Ohanian. The four-quarter random walk benchmark is nested in the AR(AIC) model, but evidently imposing the four-quarter random walk restriction matters considerably.

6. The UC-SV model imposes a unit root in inflation, so it is consistent with the Pivetta-Reis (2007) evidence that the largest AR root in inflation has been essentially one throughout the postwar sample. But the time-varying relative variances of the permanent and transitory innovation allow for persistence to change over the course of the sample and for spectral measures of persistence to decline over the sample, consistent with Cogley and Sargent (2002, 2005).

7. Koenig (2003, table 3) presented in-sample evidence that real-time markups (nonfinancial corporate GDP divided by nonfinancial corporate employee compensation), in conjunction with the unemployment rate, significantly contribute to a forecast combination regression for four-quarter CPI inflation over 1983–2001; however he did not present pseudo out-of-sample RMSEs. Two of Ang, Bekaert, and Wei's (2007) models (their PC9 and PC10) include the output gap and the labor income share, specifications similar to the Koenig's (2003), and the pseudo out-of-sample performance of these models is poor: over Ang, Bekaert, and Wei's (2007) two subsamples and four inflation measures, the RMSEs, relative to the ARMA(1,1) benchmark, range from 1.17 to 3.26. These results sug-

gest that markups are not a solution to the poor performance of Phillips curve forecasts over the post-85 samples.

8. Romer and Romer (2000) compared the performance of real-time professional inflation forecasts and found that Fed Greenbook forecasts outperform commercial forecasts (Data Resources, Inc., Blue Chip, Survey of Professional Forecasters) over the period starting 1968:M11–1980:M1 (the start date depends on the forecast source) through 1999:M11. Romer and Romer's (2000) findings do not speak directly to the inflation forecasting literature discussed here, however, because they do not analyze performance relative to a univariate benchmark, nor do they report results for post-1984 subsamples.

9. Cecchetti et. al. (2007, section 7) provided in-sample evidence that survey inflation forecasts are correlated with future trend inflation, measured using the Stock-Watson (2007) UC-SV model. Thus a different explanation of why surveys perform well is that survey inflation expectations anticipate movements in trend inflation.

10. The exceptions are rolling forecasts for the 2001–2007 sample: for CPI-core inflation using median CPI as a predictor, and for GDP inflation using either median or trimmed-mean CPI as a predictor. However, the relative RMSEs exceed one (typically, they exceed 1.15) for other inflation series, other samples, and for recursive forecasts, and we view these three exceptional cases as outliers. Most likely, the difference between our negative results for median CPI and Smith's (2004) positive results over 1990-2000 are differences in the benchmark model, which in her case is a univariate AR with exponential lag structure imposed.

References

Ang, Andrew, Geert Bekaert, and Min Wei. 2007. "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?" *Journal of Monetary Economics* 54(4): 1163–1212.

Ascari, Guido, and Emanuela Marrocu. 2003. "Forecasting Inflation: A Comparison of Linear Phillips Curve Models and Nonlinear Time Series Models." Working Paper CRENoS 200307, Centre for North South Economic Research, University of Cagliari, Italy. Available at http://veprints.unica.it/273/1/03-07.pdf.

Atkeson, Andrew, and Lee E. Ohanian. 2001. "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review* 25(1): 2–11. Available at http://www.minneapolisfed.org/research/QR/QR2511.pdf.

Bańbura, Marta, Domenico Gianonne, and Lucrezia Reichlin. 2008. "Large Bayesian VARs." Working Paper No. 966. Frankfurt: European Central Bank. Available at http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp966.pdf.

Banerjee, Anindya, and Massimiliano Marcellino. 2006. "Are There Any Reliable Leading Indicators for U.S. Inflation and GDP Growth?" *International Journal of Forecasting* 22(1): 137–151.

Barnes, Michelle L., and Giovanni P. Olivei. 2003. "Inside and Outside the Bounds: Threshold Estimates of the Phillips Curve." *New England Economic Review*: 3–18. Available at http://www.bos.frb.org/economic/neer/neer2003/neer03a.pdf.

Barsky, Robert B. 1987. "The Fisher Hypothesis and the Forecastability and Persistence of Inflation." *Journal of Monetary Economics* 19(1): 3–24.

Bates, J.M., and Clive W.J. Granger. 1969. "The Combination of Forecasts." *Operations Research Quarterly* 20: 451–468.

Berardi, Andrea. 2007. "Term Structure, Inflation, and Real Activity." Manuscript. Available at http://dse.uiver.it/berardi/tsira.pdf.

Bernanke, Ben S. 2003. "An Unwelcome Fall in Inflation?" Speech delivered at the Economics Roundtable, University of California, San Diego, July 23. Available at http://www.federalreserve.gov/boarddocs/speeches/2003/20030723/.

Bernanke, Ben S., Jean Bovian, and Piotr Eliasz. 2005. "Measuring the Effects of Monetary Policy: a Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics* 120(1): 387–422.

Boivin, Jean, and Serena Ng. 2005. "Understanding and Comparing Factor-Based Forecasts." *International Journal of Central Banking* 1(December): 117–151. Available at http://www.ijcb.org/journal/ijcb05q4a4.pdf.

Boivin, Jean, and Serena Ng. 2006. "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics* 132(1): 169–194.

Brainard, William C., and George L. Perry. 2000. "Making Policy in a Changing World." In *Economic Events, Ideas, and Policies: The 1960s and After*, ed. George L. Perry and James Tobin, 43–68. Washington, DC: Brookings Institution Press.

Brayton, Flint, John M. Roberts, and John C. Williams. 1999. "What's Happened to the Phillips Curve?" Finance and Economics Discussion Series 1999-49. Washington, DC: Board of Governors of the Federal Reserve. Available at http://www.federalreserve.gov/pubs/feds/1999/199949/199949pap.pdf.

Brave, Scott, and Jonas D.M. Fisher. 2004. "In Search of a Robust Inflation Forecast." Federal Reserve Bank of Chicago *Economic Perspectives* 28 (4): 12–31. Available at http://www.chicagofed.org/publications/economicperspectives/ep_ 4qtr 2004_part2_Brave_Fisher.pdf.

Bryan, Michael F., and Stephen G. Cecchetti. 1994. "Measuring Core Inflation." In *Monetary Policy*, ed. N. Gregory Mankiw, 195–215. Chicago: University of Chicago Press.

Bryan, Michael F., Stephen G. Cecchetti, and Rodney L. Wiggins II. 1997. "Efficient Inflation Estimation." Working Paper No. 6183. Cambridge, MA: National Bureau of Economic Research.

Camba-Mendez, Gonazlo, and Diego Rodriguez-Palenzuela. 2003. "Assessment Criteria for Output Gap Estimates." *Economic Modeling* 20(3): 528–561.

Canova, Fabio. 2007. "G-7 Inflation Forecasts: Random Walk, Phillips Curve or What Else?" *Macroeconomic Dynamics* 11(1): 1–30.

Cecchetti, Stephen G. 1995. "Inflation Indicators and Inflation Policy." In *NBER Macroeconomics Annual* 10, ed. Ben S. Bernanke and Julio J. Rotemberg, 189–219. Cambridge, MA: The MIT Press.

Cecchetti, Stephen G., Rita S. Chu, and Charles Steindel. 2000. "The Unreliability of Inflation Indicators." Federal Reserve Bank of New York *Current Issues in Economics and Finance* 6(4):1–6. Available at http://www.newyorkfed.org/research/current_issues/ci6-4.pdf.

Cecchetti, Stephen G., Peter Hooper, Bruce C. Kasman, Kermit L. Schoenholtz, and Mark W. Watson. 2007. "Understanding the Evolving Inflation Process." Report Prepared for the 2007 U.S. Monetary Policy Forum. Available at http://research.chicagogsb.edu/igm/docs/2007USMPF-Report.pdf.

"The Chicago Fed National Activity Index." Various dates. Chicago: Federal Reserve Bank of Chicago. Available at http://www.chicagofed.org/economic_research_and_data/cfnai.cfm.

Clark, Todd E., and Michael W. McCracken. 2006. "'The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence." *Journal of Money, Credit and Banking* 38(5): 1127–1148.

Clark, Todd E., and Michael W. McCracken. 2006; revised 2008. "Combining Forecasts from Nested Models." Research Working Paper 06-02. Kansas City, MO: Federal Reserve Bank of Kansas City. Available at http://www.kansascityfed.org/Publicat/Reswkpap/PDF/RWP06-02v2.pdf.

Cogley, Timothy, and Thomas J. Sargent. 2002. "Evolving Post-World War II U.S. Inflation Dynamics." In *NBER Macroeconomics Annual* 2001, ed. Ben S. Bernanke and Kenneth Rogoff, 331–388. Cambridge, MA: The MIT Press.

Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post World War II U.S." *Review of Economic Dynamics* 8(2): 262–302.

D'Agostino, Antonello, and Domenico Giannone. 2006. "Comparing Alternative Predictors Based on Large-Panel Factor Models." Working Paper Series 680. Frankfurt: European Central Bank. Available at http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp680.pdf.

Diebold, Francis X., and Roberto S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13(3): 253–263.

DeWachter, Hans, and Marco Lyrio. 2006. "Macro Factors and the Term Structure of Interest Rates." *Journal of Money, Credit, and Banking* 38(1): 119–140.

Diron, Marie, and Benoît Mojon. 2008. "Are Inflation Targets Good Inflation Forecasts?" Federal Reserve Bank of Chicago *Economic Perspectives* 32(2): 33–45. Available at http://www.chicagofed.org/publications/economic perspectives/ep_2qtr2008_part3_diron_mojon.pdf.

Dupasquier, Chantal, and Nicholas Ricketts. 1998. "Nonlinearities in the Output-Inflation Relationship." In *Price Stability, Inflation Targets and Monetary Policy*, 131–167. Ottawa: Bank of Canada. Available at http://www.bankofcanada.ca/en/conference/con97/cn97-8.pdf.

Eichmeier, Sandra, and Christina Ziegler. 2008. "How Good are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach." *Journal of Forecasting* 27(3): 237–265.

European Central Bank Inflation Persistence Network. Various Dates. Research papers available at http://www.ecb.eu/home/html/researcher_ipn_papers.en.html.

Estrella, Arturo, and Jeffrey C. Fuhrer. 2003. "Monetary Policy Shifts and the Stability of Monetary Policy Models." *Review of Economics and Statistics* 85(1): 94–104.

Fisher, Jonas D.M., ChinTe Liu, and Ruilin Zhou. 2002. "When Can We Forecast Inflation?" Federal Reserve Bank of Chicago *Economic Perspectives* 26(1): 30–42. Available at http://www.chicagofed.org/publications/economic perspectives/2002/1qepart4.pdf.

Gavin, William T., and Kevin L. Kliesen. 2008. "Forecasting Inflation and Output: Comparing Data-Rich Models with Simple Rules." Federal Reserve Bank of St. Louis *Review* 90(3, Part 1): 175–192. Available at http://research.stlouisfed.org/publications/review/08/05/GavinKliesen.pdf.

Geweke, John. 1977. "The Dynamic Factor Analysis of Economic Time Series." In *Latent Variables in Socio-Economic Models*, ed. Dennis J. Aigner and Arthur S. Goldberger, 365–383. Amsterdam: North-Holland.

Giacomini, Raffaella, and Halbert White. 2006. "Tests of Conditional Predictive Ability." *Econometrica* 74(6): 1545–1578.

Gordon, Robert J. 1982. "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment." In *Workers, Jobs and Inflation*, ed. Martin N. Baily, 89–158. Washington, DC: The Brookings Institution.

Gordon, Robert J. 1990. "U.S. Inflation, Labor's Share, and the Natural Rate of Unemployment." In *Economics of Wage Determination*, ed. Heinz Konig, 1–34. New York: Springer-Verlag.

Gordon, Robert J. 1998. "Foundations of the Goldilocks Economy: Supply Shocks and the Time-Varying NAIRU." *Brookings Papers on Economic Activity* 2: 297–333.

Hansen, Peter Reinhard. 2005. "A Test for Superior Predictive Ability." *Journal of Business and Economic Statistics* 23(4): 365–380. Available at http://pubs.amstat.org/doi/pdfplus/10.1198/073500105000000063.

Hendry, David F., and Michael P. Clements. 2004. "Pooling of forecasts." *Econometrics Journal* 7(1): 1–31.

Hendry, David F., and Kirstin Hubrich. 2007. "Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate." Paper presented at Oxford University. Conference in Honour of David F. Hendry, 23 August. Oxford: Oxford University. Available at http://www.economics.ox.ac.uk/hendryconference/Papers/Hubrich_DFHVol.pdf.

Hubrich, Kirstin. 2005. "Forecasting Euro Area Inflation: Does Aggregating Forecasts by HICP Component Improve Forecast Accuracy?" *International Journal of Forecasting* 21(1): 119–136.

Inoue, Atsushi, and Lutz Kilian. 2008. "How Useful is Bagging in Forecasting Economic Time Series? A Case Study of U.S. CPI Inflation." *Journal of the American Statistical Association* 103(482): 511–522.

Jaditz, Ted, and Chera Sayers. 1994. "Predicting Inflation." Manuscript, Bureau of Labor Statistics.

Jorion, Philippe, and Frederic Mishkin. 1991. "A Multi-Country Comparison of Term Structure Forecasts at Long Horizons." *Journal of Financial Economics* 29(1): 59–80.

König, Evan F. 2003. "Is the Markup a Useful Real-Time Predictor of Inflation?" *Economics Letters* 80(2): 261–267.

Levin, Andrew T., and Jeremy M. Piger. 2004. "Is Inflation Persistence Intrinsic in Industrial Economics?" Working Paper Series 334. Frankfurt: European Central Bank. Available at http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp334.pdf.

Lown, Cara S., and Robert W. Rich. 1997. "Is There an Inflation Puzzle?" Federal Reserve Bank of New York *Economic Policy Review* 3(4): 51–69. Available at http://www.newyorkfed.org/research/epr/97v03n4/9712lown.pdf.

Marcellino, Massimiliano. "A Linear Benchmark for Forecasting GDP Growth and Inflation?" *Journal of Forecasting* 27(4): 305–340.

Mishkin, Frederic S. 1990a. "What Does the Term Structure Tell Us About Future Inflation?" *Journal of Monetary Economics* 25(1): 77–95.

Mishkin, Frederic S. 1990b. "The Information in the Longer-Maturity Term Structure About Future Inflation." *Quarterly Journal of Economics* 105(3): 815–828.

Mishkin, Frederic S. 1991. "A Multi-Country Study of the Information in the Term Structure About Future Inflation." *Journal of International Money and Finance* 10(1): 2–22.

Mishkin, Frederic S. 2007. "Inflation Dynamics." Working Paper 13147. Cambridge, MA: National Bureau of Economic Research.

Moshiri, Saeed, and Norman Cameron. 2000. "Neural Network Versus Econometric Models in Forecasting Inflation." *Journal of Forecasting* 19(3): 201–217.

Nelson, Charles R., and G. William Schwert. 1977. "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis that the Real Rate of Interest Is Constant." *American Economic Review* 67(3): 478–486.

Orphanides, Athanasios, and Simon van Norden. 2005. "The Reliability of Inflation Forecast Based on Output Gap Estimates in Real Time." *Journal of Money, Credit and Banking* 37(3): 583–600.

Pivetta, Frederic, and Ricardo Reis. 2007. "The Persistence of Inflation in the United States." *Journal of Economic Dynamics and Control* 31(4): 1326–1358.

Primiceri, Giorgio E. 2006. "Why Inflation Rose and Fell: Policy-makers' Beliefs and U.S. Postwar Stabilization Policy." *The Quarterly Journal of Economics* 121(3): 867–901.

Reis, Ricardo, and Mark W. Watson. 2007. "Relative Goods' Prices and Pure Inflation." Working Paper No. 13615. Cambridge, MA: National Bureau of Economic Research.

Roberts, John M. 1995. "New Keynesian Economics and the Phillips Curve." *Journal of Money, Credit and Banking* 27(4): 975–984.

Roberts, John M. 2004. "Monetary Policy and Inflation Dynamics." *International Journal of Central Banking* 2(3): 193–230. Available at http://www.ijcb.org/journal/ijcb06q3a6.pdf.

Romer, Christina D., and David H. Romer. 2000. "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review* 90(3): 429–457.

Rossi, Barbara, and Tatevik Sekhposyan. 2007. "Has Models' Forecasting Performance for U.S. Output Growth and Inflation Changed Over Time, and When?." Manuscript, Duke University. Available at http://www.econ.duke.edu/-brossi.

Rünstler, Gerhard. 2002. "The Information Content of Real-Time Output Gap Estimates: An Application to the Euro Area." Working Paper Series 182. Frankfurt: European Central Bank.

Sargent, Thomas J., and Christopher A. Sims. 1977. "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory." Working Paper No. 55. Minneapolis: Federal Reserve Bank of Minneapolis. Available at http://www.minneapolisfed.org/research/WP/WP55.pdf.

Sims, Christopher A. 2002. "The Role of Models and Probabilities in the Monetary Policy Process." *Brookings Papers on Economic Activity* 2: 1–40.

Sims, Christopher A., and Tao Zha. 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* 96(1): 54–81.

Smith, Julie K. 2004. "Weighted Median Inflation: Is This Core Inflation?" *Journal of Money, Credit and Banking* 36(2): 253–263.

Staiger, Douglas, James H. Stock, and Mark W. Watson. 1997. "The NAIRU, Unemployment, and Monetary Policy." *Journal of Economic Perspectives* 11(1): 33–51.

Stock, James H., and Mark W. Watson. 1999. "Forecasting Inflation." *Journal of Monetary Economics* 44(2): 293–335.

Stock, James H., and Mark W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97(460): 1167–1179.

Stock, James H., and Mark W. Watson. 2003. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature* 41(3): 788–829.

Stock, James H., and Mark W. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39(1): 3–34.

Stockton, David J., and James E. Glassman. 1987. "An Evaluation of the Forecast Performance of Alternative Models of Inflation." *The Review of Economics and Statistics* 69(1): 108–117.

Taylor, John B. 2000. "Low Inflation, Pass-Through, and the Pricing Power of Firms." *European Economic Review* 44(7): 1389–1408.

Timmermann, Allan. 2006. "Forecast Combinations." In *Handbook of Economic Forecasting*, Vol. 1, ed. Graham Elliott, Clive W.J. Granger, and Allan Timmermann, 135–196. Amsterdam: North-Holland.

Tkacz, Greg. 2000. "Non-Parametric and Neural Network Models of Inflation Changes." Working Paper 2000-7. Ottawa: Bank of Canada. Available at http://www.bankofcanada.ca/en/res/wp/2000/wp00-7.pdf.

Watson, Mark W. 2003. "Macroeconomic Forecasting Using Many Predictors." In *Advances in Econometrics: Theory and Applications*, Eighth World Congress, ed. Mathias Dewatripont, Lars Peter Hansen, and Stephen Turnovsky, 87–115. New York: Cambridge University Press.

Webb, Roy H., and Tazewell S. Rowe. 1995. "An Index of Leading Indicators for Inflation." Federal Reserve Bank of Richmond *Economic Quarterly* 81(2): 75–96. Available at http://www.richmondfed.org/publications/research/economic_quarterly/1995/spring/pdf/webb.pdf.

West, Kenneth D. 1996. "Asymptotic Inference about Predictive Accuracy." *Econometrica* 64(5): 1067–1084.

Wright, Jonathan. 2003. "Forecasting U.S. Inflation by Bayesian Model Averaging." International Finance Discussion Paper No. 780. Washington, DC: Board of Governors of the Federal Reserve System. Available at http://www.federalreserve.gov/pubs/ifdp/2003/780/ifdp780.pdf.

Data Appendix

The definitions and sources of the series used in this analysis are summarized in the following table. The "trans" column indicates the transformation applied to the series: logarithm (ln), first difference of logarithm ($(1-L)\ln$), accumulation ($(1-L)^{-1}$), or no transformation (level). When the original series is monthly, quarterly data are constructed as the average of the monthly values in the quarter before any other transformation. Sources are Federal Reserve Bank of St. Louis FRED database (F), the Bureau of Economic Analysis (BEA), and other Federal Reserve banks as indicated.

Short name	Trans	Definition	Mnemonic (Source)
		Inflation series	
CPI-all	(1-L)ln	CPI, all items	CPIAUCSL (F)
CPI-core	(1-L)ln	CPI less food and energy	CPILFESL (F)
PCE-all	(1-L)ln	PCE deflator, all items	PCECTPI (F)
PCE-core	(1-L)ln	PCE deflator, less food and energy	JCXFE (F)
GDP deflator	(1-L)ln	GDP deflator	GDPCTPI (F)
		Predictors	
UR	level	Unemployment rate, total civilian 16+	UNRATE (F)
GDP	ln	Real GDP	GDPC96 (F)
IP	ln	Index of Industrial Production (total)	INDPRO (F)
EMP	ln	Nonagricultural civilian employment (total)	PAYEMS (F)
CapU	level	Capacity utilization rate	TCU (F)
HPerm	ln	Housing permits (starts)	PERMIT (F)
CFNAI	$(1–L)^{-1}$	Chicago Fed National Activity Index (accumulated)	FRB-Chicago
UR-5wk	level	Unemployment rate for unemployed < 5 week	UEMPLT5(F) / CLF160V(F)
AHE	(1-L)ln	Average hourly earnings	AHETPI (F)
Real AHE	(1-L)ln	real average hourly earnings	AHETPI (F)/ GDPCTPI (F)
Labor Share	ln	labor share	AHETPI (F)/ GDPCTPI (F)
CPI-Median	level	Cleveland Fed median CPI inflation "Original" CPI-Median through 2007:M7; "Revised" CPI-Median after 2007:M7)	FRB-Cleveland
CPI-TrMn	level	Cleveland Fed trimmed mean CPI inflation ("Original" CPI-Trimmed Mean through 2007:M7; "Revised" CPI-Trimmed Mean after 2007:M7)	FRB-Cleveland
ExRate	level	trade-weighted exchange rate	TWEXMMTH (F)
TB_sp	level	1 Year Treasury bond rate minus 3 Month Treasury bill rate (at annual rate)	Fed Board of Governors
RPFE	(1–L)ln	Relative Price of Food and Energy	PCECTPI (F)/ JCXFE (F)
RPImp	(1–L)ln	Relative Price of Imports	B021RG3(BEA)/ GDPCTPI(F)
Price Control Variable 1	level	0.8 for 1971:Q3≤ t≤1972:Q2, 0 otherwise	Gordon (1982)
Price Control Variable 2	level	–0.4 for t = 1974:Q2 or 1975:Q1, –1.6 for 1974:Q3≤ t≤1974:Q4, 0 otherwise.	Gordon (1982)

Comments on "Phillips Curve Inflation Forecasts" by James H. Stock and Mark W. Watson

Adrian Rodney Pagan

Stock and Watson provide a thorough (one might say exhaustive) review of the forecasting performance of many inflation models. These include models with some economics in them and others that are purely statistical. Overall, the best of the statistical models seems to be their unobserved components-stochastic volatility (UC-SV) model. This is a two-component model of the form

$$\pi_t = z_t^* + v_t$$

and

$$z_t^* = z_{t-1}^* + u_t .$$

As written, z_t^* represents a permanent component to inflation and v_t is a transitory one. The shocks to the components are taken as having time-varying stochastic volatilities (SV) σ_{jt}^2 of the form $\ln\sigma_{jt}^2 = \ln\sigma_{jt-1}^2 + \eta_{jt}$. Consequently these have a unit root. The variances of the shock terms η_{jt} are equal. This would mean that the true volatilities would also be equal if the initial conditions were the same. The estimated ones can however vary as they depend upon the data.

Statistical models are generally judged by how well they fit and forecast over limited sample sizes and horizons. That is probably just as well here since it would be hard to think about targeting an inflation rate that really behaved like the UC-SV model, since it has inflation following a unit root and with the variances of v_t and u_t being unbounded, meaning there is no second moment for the change in inflation. If you try to simulate a process like UC-SV, it blows up very quickly.

I guess I am rather doubtful about whether I want to use a model like this, even if it produces good forecasts, as it would be hard to believe that we could keep inflation in a target range for very long if it behaved in such a way. It is only over short periods that this model makes much sense. Given this reality, I wonder why Stock and Watson set up the UC model with a permanent shock. Indeed, one could have instead chosen a very persistent process such as $z_t^* = .99z_{t-1}^* + u_t$. Since putting any number on the degree of persistence is arbitrary, one might as well use something like .99, because that ties in better with what we hope is the nature of the inflation process. In doing so, no extra coefficients would be estimated and it seems highly likely that over short forecast horizons the forecasts using both sets of parameters would be very close. It might be different if we looked at an eight-period horizon, since most central banks forecast both one and two years ahead.

Stock and Watson conclude that there are periods of time when the UC-SV model can be beaten by models featuring economic variables, principally when there are large departures from the NAIRU or when one is in an extreme recession. But for most of the time the economic variables don't contribute much to forecasts. Of course there are good reasons why we still believe that economic variables are influential, even if we cannot detect a precise role for these variables in forecasts. It is a well-known fact that relatively simple statistical models win forecasting competitions. But in a world in which we are increasingly forced to explain policy actions, any forecasts underlying them, and the risks associated with those forecasts, statistical model forecasts are clearly of limited use. In practice many central banks use a Phillips-curve type equation to give a central forecast—and even judgmental forecasts often have this as a base—relegating the statistical model forecasts to the role of checking and "tweaking" the central forecast. It is hard to imagine any central banker not putting some faith in the role of excess demand in accounting for inflation outcomes, even if measures of excess demand do little for forecasting inflation a year ahead. There are many reasons why we might see this play a role in forecasting inflation. The excess demand may need to be sustained for a long time, many measures of it are exceedingly volatile, inflation itself can have a lot of noise, and it may only be when a threshold is exceeded that there are substantial effects on inflation. These are all hard to measure precisely in a model given the length of the data sets we typically have to work with. But at some point, economic variables need to enter the forecast—otherwise it is not going to be easy to explain any actions you take as a consequence of the forecasts.

So let us remind ourselves why the UC-SV model might win the Stock and Watson forecasting competition. To do this, assume that there is no SV in the UC model and that the variances of the two shocks u_t and v_t are in a fixed ratio q. Then, it is well known that the forecast from the UC model is (provided $E_t(v_{t+1}) = 0$)

$$E_t\pi_{t+1} = (1-\phi)\sum_{j=0}^{\infty}\phi^j \pi_{t-j}.$$

The Kalman predictor is used on the UC model to find that ϕ solves $\phi + q\phi^2 - 1 = 0$, q can be found from $q = -2 - \frac{1}{\rho_1}$, and ρ_1 is the first-order serial correlation coefficient of $\Delta\pi_t$. This is the exponentially weighted moving average (EWMA) forecasting formula that is widely used in industry for forecasting the level of product demand. So perhaps it is not surprising that it produces a good forecast for inflation. This formula has also been used recently in the financial literature to forecast series that are random walks in which the drift term changes over time; in other words, the EWMA formula has some robustness to breaks in those series. So it may be a good vehicle for inflation forecasting as well.

To develop this theme further, note that $\rho_1 = \frac{\alpha}{1+\alpha^2}$, where α is the moving average coefficient in the autoregressive integrated moving average (ARIMA) representation of the UC model, or $\Delta\pi_t = \varepsilon_t + \alpha\varepsilon_{t-1}$. Stock and Watson note that α has been varying a good deal over U.S. history and, recently, $\alpha \cong -.85$, and so $\rho_1 = -.493$, implying that q is close to zero (and ϕ is close to unity), meaning there is little weight placed on past inflation. Indeed a large negative value of α is really consistent with inflation not having a unit root, and the way this shows up in the UC model is for the variance of the transitory shocks to become large relative to the permanent ones.

Now one reason why the EWMA forecasting formula has been popular is that it represents a simple forecasting mechanism that is relatively robust to structural change, provided one modifies q at different times. A plot of the inflation rate suggests that such changes have occurred, and there is an extensive literature maintaining that such breaks have occurred in many countries in the past three decades, as well as in the United States. The key question then becomes how to vary q in response to such developments, and this is what the SV part of Stock and Watson's UC-SV model does, since the estimated relative volatilities can change over time, leading to a change in q.

So this leads us to ask what the implications of the Stock and Watson paper are for forecasting. This may not be a fair question since their brief seems to have been to ask if Phillips curve-type economic variables are useful for forecasting rather than to ask what is the best forecasting method. So I am possibly being unfair when I ask if their UC-SV model is the best forecasting mechanism, but I think their results are sufficiently striking for me to make some comments on this.

As mentioned above, in Stock and Watson's case q adapts to the data to account for breaks through the relative size of the stochastic volatilities. Are there other ways of doing this? Pesaran and Timmermann (2007) point out that we need to detect when a break in the inflation process took place and also the size of the break; stated differently, we need to know when to vary q and by how much. There has been much research on methods to detect a break in inflation but less on determining the size. However there is now an emerging literature on techniques to gauge the size of the break. Pesaran and Timmermann (2007) have proposed the idea that one should average forecasts not across models (Stock and Watson show that this does not give much advantage) but across the windows over which parameter estimation is performed prior to making the forecast. Pesaran and Timmermann demonstrate that this can yield improvements in forecasting a random walk process in the presence of breaks. Pesaran and Pick (2008) show that there are theoretical reasons to expect that this procedure will improve forecasts in the presence of breaks. They also look at EWMA forecasts for different q values and then average the forecasts from these values. An advantage of focusing upon the EWMA approach is that no judgment is being made about the nature of the inflation process (one still uses a weighted average of inflation rates). If the inflation series was white noise, then one would simply put $q = 0$. So I think it would be interesting to compare the forecasts from this methodology with those from the UC-SV model.

Now let us think about introducing economic information into the forecasts. Traditionally this involved replacing z_t^* with some functions of lagged inflation, unemployment, supply-side effects, deviations of this from the NAIRU, expectations, and so on. In the U.S. literature this is often referred to as Gordon's triangle model, although models like it have been present in many countries since the 1970s. Whatever individual information is introduced needs to be combined together to produce a persistent component. The flaws in the approach are the need to estimate the parameters in any such relation when providing forecasts of these covariates. Consequently, it is possible that we will do worse than a model that ignores the effects, even if we believed that the economics in such a model tells us something about what might have driven an observed inflation path. That will almost certainly happen if the parameters are imprecisely determined, and one would have to say that this is indeed true of most Phillips curve models. Moreover many of the variables added into the relation can change quite dramatically as a result of data revisions. So even if the forecasts are good with data that has been finally revised, the need to use real-time data may result in the forecasts being quite poor—see Robinson, Stone, and van Zyl (2003) for an Australian example. Since Stock and Watson did not use real-time data it would seem that the Phillips curve-based forecasts they report would most likely be better than these would be in real time. It is only with extreme movements in the determining variables—very high unemployment relative to the NAIRU or expectations relative to, say, the target—that we can observe big enough effects on inflation to offset these difficulties. However it should be noted that it may be possible to use economic information to reliably signal the direction of change in inflation, as found in Robinson, Stone, and van Zyl (2003) for Australia and in Fisher, Liu, and Zhou (2002) for the United States, and in many contexts this might well be sufficient.

Finally, we might just make ϕ a function of some economic variables and so change the exponential weights. To get some idea of whether this would work we need a series on ϕ_t. I fitted an AR(1) to the change in inflation using a ten-year rolling horizon to get an estimate of ρ_1 that changes over time. From that I got values for q_t and ϕ_t. Figure 3.18 shows

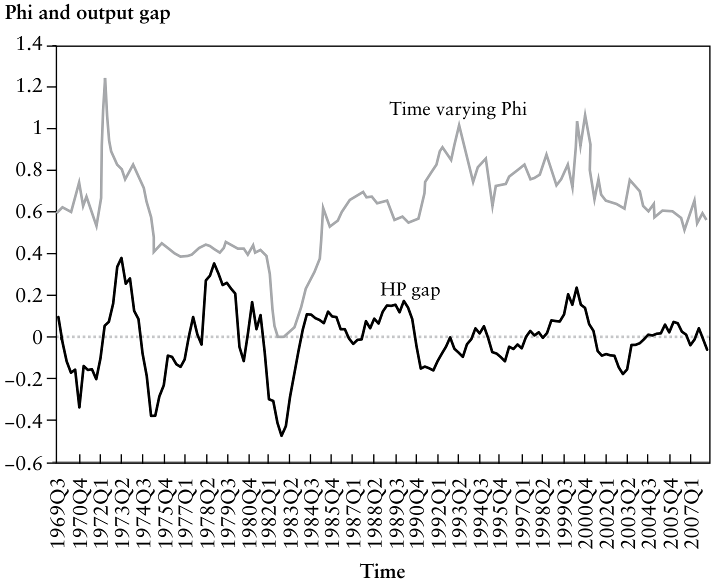


Figure 3.18
Time-Varying Phi and the Hodrick-Prescott-Filtered Output Gap

the series on ϕ_t obtained with this method. Note that there are some missing observations in the 1982:Q3–1983:Q2 period since the estimate of ρ_1 during that time was positive, which is not admissible if the UC model is correct. Consequently, I have set this particular period's observations to zero to match the movements that were evident before and after that period. The graph is best at identifying changes in ϕ_t rather than the precise values of it, but it shows that the weighting factor on past data to be used when forecasting has varied significantly over history. In the 1982 recession the close-to-zero weights pointed to ignoring the past history of inflation and indicated that some other information needed to be used. The output gap (found with the Hodrick-Prescott filter) in figure 3.18 has a positive correlation with ϕ_t, and so measures of excess demand are likely to be useful for forecasting. This finding reinforces Stock and Watson's conclusion that the Phillips curve was useful in forecasting inflation in the first half of the 1980s.

References

Fisher, Jonas D.M., Chin Te Liu, and Ruilin Zhou. 2002. "When Can We Forecast Inflation?" *Federal Reserve Bank of Chicago Economic Perspectives* Q1: 30–42. Available at http://www.chicagofed.org/publications/economic perspectives/2002/1qepart4.pdf.

Pesaran, M. Hashem, and Allan Timmermann. 2007. "Selection of Estimation Window in the Presence of Breaks." *Journal of Econometrics* 137(1): 134–161.

Pesaran, M. Hashem, and Andreas Pick. 2008. "Forecasting Random Walks Under Drift Instability." Cambridge Working Papers in Economics No. 0814. Faculty of Economics, University of Cambridge. Available at http://www.econ.cam.ac.uk/dae/repec/cam/pdf/cwpe0814.pdf.

Robinson, Tim, Andrew Stone, and Marileze van Zyl. 2003. "The Real-Time Forecasting Performance of Phillips Curves." Economic Research Department Discussion Paper 2003–12. Sydney: Reserve Bank of Australia. Available at http://www.rba.gov.au/rdp/RDP2003-12.pdf.

Comments on "Phillips Curve Inflation Forecasts" by James H. Stock and Mark W. Watson

Lucrezia Reichlin

1. Is the Phillips Curve Dead?

Stock and Watson present convincing evidence on the Phillips curve's lack of predictive power over the last fifteen years—meaning the inability of models based on the Phillips relationship between inflation and unemployment to predict beyond a naïve benchmark such as the random walk.

The analysis is very convincing: it is conducted systematically on the basis of different specifications and definitions of real indicators of economic activity and is based on a simulated out-of-sample exercise. This is the right methodology to evaluate the robustness of the Phillips relation since, unlike in regression analysis, the evaluation results are not only valid under the assumption of the correct specification of the model.

The result is perhaps disturbing to the macroeconomic profession which has spent so much time formulating and discussing micro-foundations for the Phillips curve. Has the profession wasted its time trying to explain a relationship that had in fact disappeared? More constructively, what can we learn about our macroeconomic models from this result?

In this discussion I will bring some complementary evidence to the authors' results, focusing not only on the predictive ability of the Phillips curve relationship but also on the predictability of inflation and real activity in general. The evidence I will present suggests a decline of relative predictability not only for inflation, but also for real economic activity and for statistical as well as institutional models such as the Greenbook forecasts prepared by the staff of the Federal Reserve Board.

I will then ask whether the decline in predictability can be explained by a change in the structure of the covariance between different variables in the economy, or if this decline is the result of the decline of the variability of exogenous shocks. To analyze this question I will propose a quantitative exercise.

Clearly, to the extent that evidence points to a change in the covariances, the decline in predictability should be attributed to changes in policy or structural features of the economy, occurring around the mid-1980s. In that case the way forward must be to build on these results and try to understand the breakdown of the Phillips curve as the endogenous result of changes in structural and policy parameters in structural models.

2. Decreasing Relative Predictability

In what follows I will focus on relative predictability as defined by the expression:

$$RP_{it} = \frac{\hat{E}[X_{it+h|t}^{model} - X_{it+h}^{2}]^2}{\hat{E}[X_{it+h|t}^{naive} - X_{it+h}^{2}]^2}.$$

Here, relative predictability is defined as the predictive ability of a given model relative to the prediction based on a random walk model and measured in terms of mean squared errors.

Recent literature has pointed to a decrease in the relative predictability of inflation. The evidence is accurately surveyed by Stock and Watson's present paper and I have little to add.

Perhaps the best way to understand the evidence is to inspect figure 3.19, which plots the GDP-deflator inflation rate (solid-diamond line) against the naïve (random walk) forecast (dashed line) and the Greenbook forecast (solid line) since 1970; the shaded areas indicate recession episodes as defined by the National Bureau of Economic Research (NBER).

The picture illustrates two points. First, since the early 1990s, when Stock and Watson's sample starts, the naïve predictor becomes more accurate due to the decline in inflation volatility. Second, for the same sample, the Greenbook and the naïve forecasts are very similar—no clear

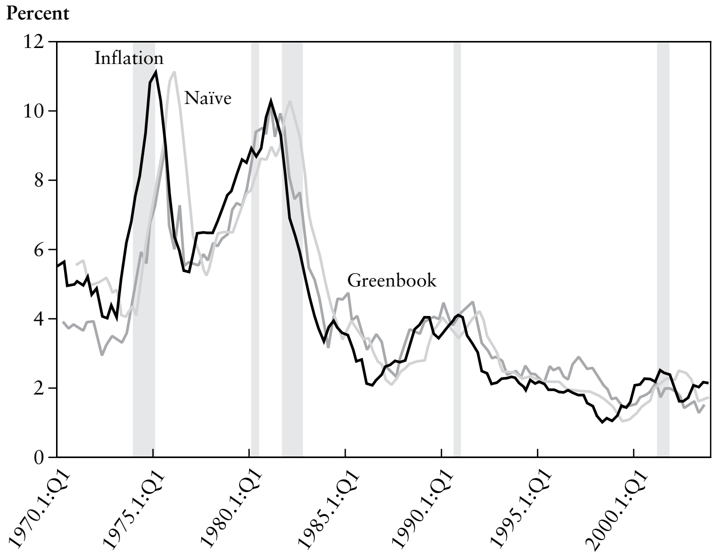


Figure 3.19
Greenbook One–Year–Ahead Forecasts of GDP Deflator Annual Inflation
Source: Author's calculations.

advantage seems to be obtained by a sophisticated forecast such as that produced by the staff of the Federal Reserve Board.

The analysis by Stock and Watson shows that the same is true for predictive equations based on the Phillips curve. The same authors have pointed out in previous work (Stock and Watson 2007) that the result applies no matter what variables are considered.

In work focusing on complex models based on a large number of predictors, nominal and real, De Mol, Giannone, and Reichlin (2008) show that principal component regression (PC), ridge regression (Ridge) and variable selection algorithms (Lasso) all produce forecasts with no relative advantage with respect to the random walk for that sample. These are relatively complex statistical models which perform very well out-of-sample until the mid-1980s. Figure 3.20 shows the forecast for the annual rate of change of Consumer Price Index (CPI) inflation since 1970 (again, the shaded areas indicate NBER-dated recessions). Clearly, in the last 20 years, the sophisticated models have not outperformed the

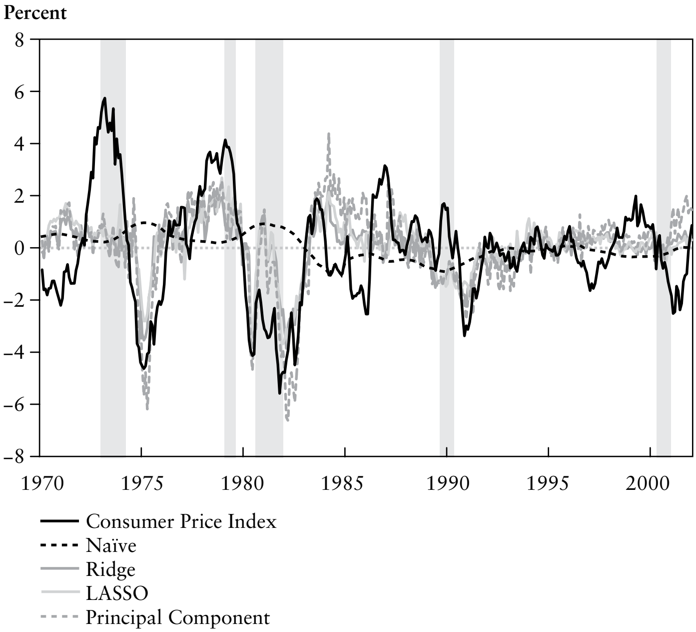


Figure 3.20
One-Year-Ahead December Forecasts of Yearly Changes of Consumer Price Index Annual Inflation
Source: De Mol, Giannone, and Reichlin (2008).

naïve model. Notice that the Lasso, PC, and Ridge forecasts are highly correlated throughout the sample and hardly distinguishable from one another.

But these results are not only true for inflation. These results remain valid for real economic activity as well, as pointed out by D'Agostino, Giannone, and Surico (2006) and by De Mol, Giannone, and Reichlin (2008). In particular, D'Agostino, Giannone, and Surico (2006) emphasize that the interest rates term structure, which was a good predictor until the mid-1980s, has failed ever since. This is an interesting result since the term structure is typically thought of being a forward-looking variable, capturing expectations of future economic activity.

For real activity, figures 3.21 and 3.22 show similar features to those described for inflation. Figure 3.21 plots the annual growth rate of GDP (solid-diamond line) since 1970, the forecast based on the Greenbook

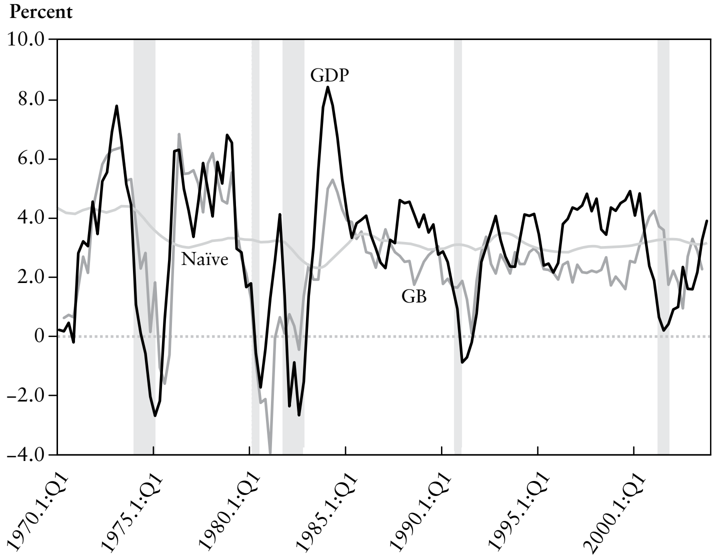


Figure 3.21
Greenbook One–Year–Ahead Forecasts for GDP Annual Growth
Source: Author's calculations.

(solid line) and the random walk (naïve) model (dashed line). Here the random walk is a model in which growth in the next period is equal to the average growth in the previous ten years.

Figure 3.22, from De Mol, Giannone, and Reichlin (2008), reports the annual change of industrial production, the naïve forecast (Naïve) again defined as a model in which growth in the next period is equal to the average growth in the previous ten years, the principal component forecast (PC), the forecast based on ridge regression (Ridge) and variable selection (Lasso) algorithms, as also shown in figure 3.20.

What I conclude from this evidence is that the Phillips curve is not the only predictive relationship that has broken down in the last 20 years. In general, we have experienced a failure of models to predict beyond a naïve benchmark, for both inflation and output.

How can we interpret this evidence? One conjecture is that the last 20 years have been a lucky period with very moderate volatility of exogenous shocks to the economy. Low volatility in exogenous shocks has

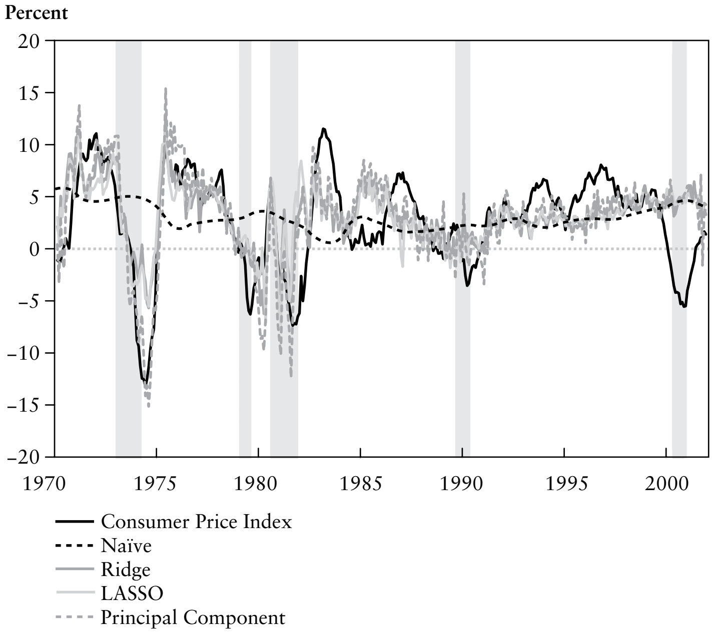


Figure 3.22
One–Year–Ahead December Forecasts for Annual Growth of Industrial Production
Source: De Mol, Giannone, and Reichlin (2008).

implied low volatility of observable variables, and in these circumstances it is not surprising that the random walk model has performed relatively well.

Alternatively, low volatility of output and inflation can be attributed to a change in the structure of the economy or to changes in policy, as extensively discussed in the literature about the Great Moderation (for a review of that debate and some new results, see Giannone, Lenza, and Reichlin 2008).

Actually, as Giannone, Lenza, and Reichlin (2008) have pointed out, decreases in inflation and output volatility combined with a decrease in the relative predictability of statistical bivariate and multivariate models and institutional models can only be explained by a change in the covariance structure of the data. Moreover, given the predictive failure of models based on many variables, including financial variables, the change must not only be in the covariance between inflation and real activity but, more generally, in the covariance between financial variables, inflation, and real activity. Finally, as we have shown in Giannone, Lenza, and Reichlin (2008), these changes can best be identified using models containing all relevant macroeconomic variables and might be lost in small models due to problems arising from omitted variables.

3. Shocks or Propagation? A Counterfactual Exercise

In this section I will ask what is the fraction of the decline of relative predictability in output and inflation that can be attributed to a decline in the variability of exogenous shocks, and what is that fraction that can be attributed to a change in the parameters of a model that include several macroeconomic indicators—among these prices, real activity, labor market, and monetary and financial indicators.

I will consider a vector autoregression (VAR) model, including 19 quarterly variables, all of which are typically used in macroeconomic models: GDP, the GDP deflator, the federal funds rate, commodity prices, consumer prices, consumption, investment, change in inventories, the producer price index, interest rates at one-, five-, and ten-year horizons, hours worked, hourly compensation, capacity utilization, stock prices, M2, total reserves, and the unemployment rate. This is the same model estimated in Giannone, Lenza, and Reichlin (2008). I refer to that paper for details on estimation and exact definitions and sources.

The VAR is estimated for two subsamples: 1959:Q1 to 1983:Q4, and 1984:Q1 to 2007:Q1.[1] Since the model is quite large and I face an issue of over-fitting, I follow Banbura, Giannone, and Reichlin (2008) and use Bayesian shrinkage. In practice I use a Litterman (random walk) prior whose tightness is set so that the in-sample fit of the interest rate equation in the large VAR models is fixed at the level achieved by a simple four-variable monetary VAR. This choice is grounded on the evidence that U.S. short-term interest rates are well described by linear functions of inflation and real activity—Taylor rules (on this point see Giannone, Lenza, and Reichlin 2008).

The models are:

$$\Delta X_t = A_{pre84}(L)\Delta X_{t-1} + e_{pre84,t} \qquad e_{pre84,t} \sim WN(0,\Sigma^{pre84}),$$

$$\Delta X_t = A_{post84}(L)\Delta X_{t-1} + e_{post84,t} \qquad e_{post84,t} \sim WN(0,\Sigma^{post84}).$$

The counterfactual exercise consists in simulating the shocks, assuming that their covariance matrix has remained unchanged at the level of the pre-1984 sample estimates ($\hat{\Sigma}_{pre84}$) and feeding them through the propagation mechanism estimated for the post-1984 sample ($\hat{A}_{post84}(L)$). Specifically, we consider the following counterfactual processes:

$$\Delta X_t^* = \hat{A}_{post84}(L)\Delta X_{t-1}^* + e_{pre84,t}^*, \quad e_{pre84,t}^* \sim WN(0,\hat{\Sigma}_{pre84}).$$

If the counterfactual relative predictability of GDP (or inflation) is the same as the actual standard deviation observed in the post-1984 sample, then this should indicate that the change of propagation mechanisms fully explains its decline. The change in shocks plays a role if, instead, the counterfactual decline is smaller than observed.

Reported in table 3.6 below, the results are unambiguous: for both inflation and GDP, the change in propagation explains all the decline in variance and all the decline in predictability.

Table 3.6
Counterfactual Volatility and Relative Predictability

		Std. Deviation		Predictability	
Coefficients	Shocks	GDP growth	Inflation	GDP growth	Inflation
		Observed			
Pre-1984	Pre-1984	2.68	2.66	0.18	0.12
Post-1984	Post-1984	1.28	0.75	0.36	0.31
		Counterfactual			
Post-1984	Pre-1984	1.30	0.69	0.47	0.33

4. Conclusion

In the last 20 years, as Stock and Watson convincingly show, the relative predictability of the Phillips curve has broken down. My discussion points out that in the same period, we have experienced a decline in the relative predictability of inflation and real activity in general. The empirical analysis I proposed suggests that these changes are attributable to changes in the multivariate covariance structure of the data.

This conclusion tells us that one direction for future research should be to study predictability as a function of the deep parameters of structural models, characterizing either structural features of the model or policy behavior. In particular, the result first presented by D'Agostino, Giannone, and Surico (2006) showing that, since the mid-1980s, the spread between short- and long-term interest rates has lost its predictive power for real activity, suggests that an important factor for declining predictability might have been changes in monetary policy which, by anchoring expectations, have broken down the predictive relation between forward-looking variables and real activity. Clearly more work is needed to explore these mechanisms.

Note

1. The models are estimated with data in log-levels except for interest rates, capacity utilization, unemployment rates, and changes in inventories, for which we do not take logarithms.

References

Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin. 2008. "Large Bayesian VARs." Working Paper No. 966. Frankfurt: European Central Bank. Available at http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp966.pdf.

D'Agostino, Antonello, Domenico Giannone, and Paolo Surico. 2006. "(Un) Predictability and Macroeconomic Stability." Working Paper Series 605. Frankfurt: European Central Bank. Available at http://www.ecb.int/pub/pdf/scpwps/ecbwp605.pdf.

De Mol, Christine, Domenico Giannone, and Lucrezia Reichlin. 2008. "Forecasting Using a Large Number of Predictors: Is Bayesian Shrinkage a Valid Alternative to Principal Components?" *Journal of Econometrics* 146(2): 318–328.

Giannone, Domenico, Michele Lenza, and Lucrezia Reichlin. 2008. "Explaining the Great Moderation: It is Not the Shocks." *Journal of the European Economic Association* 6(2–3): 621–633.

Stock, James H., and Mark W. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39(1): 3–34.