Has Inflation Become Harder to Forecast?

James H. Stock
Department of Economics, Harvard University
and the National Bureau of Economic Research

and

Mark W. Watson*
Woodrow Wilson School and Department of Economics, Princeton University
and the National Bureau of Economic Research

September 2005

*Prepared for the conference, “Quantitative Evidence on Price Determination,” Board of Governors of the Federal Reserve Board, September 29-30, Washington DC.
We thank Jonas Fisher for bringing several of the issues discussed in this paper to our attention in a 1999 conversation, and Luca Benati for (more recent) helpful suggestions. This research was funded in part by NSF grant SBR-0214131.

1. Introduction

Forecasts of the rate of price inflation play a central role in the formulation of monetary policy, and forecasting inflation is a key job for economists at the Federal Reserve Board. This paper examines whether this task has become harder and, to the extent that it has, what have been the changes in the joint time series properties of inflation and its main predictors that have made it so.

As it happens, inflation has both become harder and easier to forecast, depending on one's point of view. On the one hand, inflation – like many other macroeconomic time series – has become much less volatile, so the root mean squared error of even naïve or relatively poor forecasts had declined since the mid-1980s. In this sense, inflation has become easier to forecast: the risk of inflation forecasts, as measured by mean squared forecast errors (MSFE), has fallen. On the other hand, the relative improvement of standard multivariate forecasting models, such as the backwards-looking Phillips curve, over a univariate benchmark has been smaller in percentage terms since the mid-1980s than before. This point was forcefully made by Atkeson and Ohanian (2001) (henceforth, AO), who found that backwards-looking Phillips curve forecasts were inferior to a naïve forecast of average twelve-month inflation by its average rate over the previous twelve months. In this sense, inflation has become harder to forecast, at least, it has become much more difficult for an inflation forecaster to provide value added beyond a univariate model. But what are the changes in the time series processes of inflation and its predictors that produced these changes?

This paper proposes a parsimonious model of the changes in the univariate process for postwar U.S. quarterly inflation, in which inflation is represented as the sum of two components, a permanent stochastic trend component and a serially uncorrelated transitory component. Since the mid 1950s, there have been large changes in the magnitude of the permanent component of inflation – specifically, in the variance of the permanent disturbance – whereas the magnitude of the transitory component has been essentially constant. According to our estimates, the size of the permanent component was moderate from the mid 1950s through approximately 1970, it was large during the 1970s through 1983, it declined sharply in the mid 1980s to its value of the 1960s, and since 1990 it declined further. Currently, the variance of the permanent disturbance is estimated to be at a record low since 1954.

The time-varying trend-cycle model implies a time-varying first order integrated moving average (IMA(1,1)) model for inflation, in which the magnitude of the MA coefficient varies inversely with the ratio of the permanent to the transitory disturbance variance. Accordingly, the MA coefficient for inflation was small (approximately .25) during the 1970s but subsequently increased (the coefficient is .65 for the 1984-2004 period).

The time-varying trend-cycle model of the univariate inflation process succinctly explains the main features of the historical performance of univariate inflation forecasts. During the 1970s the inflation process was well approximated by a low order autoregression (AR), but in the mid 1980s the coefficients of that autoregressions changed and, even if these changes are taken into account, the low order autoregression became a less accurate approximation to the inflation process since 1984. The changing AR coefficients and the deterioration of the low-order AR approximation accounts for the relatively poor performance of recursive and rolling AR forecasts in the 1984-2004 sample. Moreover, it turns out that the AO year-upon-year forecast, represented as a linear combination of past inflation, is close to the optimal linear combination implied by the post-1984 IMA model at the four-quarter horizon, although this is not the case at shorter horizons for the post-1984 period or any horizon for the pre-1984 period, cases in which the AO forecasts perform relatively poorly.

This time-varying trend-cycle model also explains the excellent recent forecasting performance of an IMA model published by Nelson and Schwert (1977), which they estimated using data from 1953 to 1971. During the 1970s and early 1980s, the variance of the permanent component was an order of magnitude larger than it was in the 1950s and 1960s, and the Nelson-Schwert (1977) model ceased to be a good approximation. During the 1980s and 1990s, however, the size of the permanent component fell back to its earlier levels, and the Nelson-Schwert (1977) model was again a good approximation.

The time-varying trend-cycle model also suggests a strategy for real-time univariate forecast. Currently the Nelson-Schwert (1977) forecast is performing very well, and the AO forecast is performing nearly as well, at least at long horizons. If the importance of the permanent component were to change again, as it has in the past, the performance of these forecasts would deteriorate. The pseudo-out-of-sample results suggest that two approaches to time-varying trend-cycle models could be effective in the face of such changes, an unobserved components model with stochastic volatility, implemented using a nonGaussian filter, and an IMA(1,1) model with moving average coefficient estimated using a ten-year rolling window of past observations. The latter approach is simpler, but adapts to changing coefficients less quickly than, the former.

The changing univariate inflation dynamics also help to explain the dramatic breakdown of recursive and rolling autoregressive distributed lag (ADL) inflation forecasts based on an activity measure. One reason for the deterioration in the performance of the ADL activity-based forecasts is that the variance of the activity measures has decreased (this is the "Great Moderation"), so in a sum-of-squares sense their predictive content, assuming no changes in coefficients, has declined. But the coefficients of the ADL models have also changed, in two ways. First, because the ADL forecasts generalize a univariate autoregression, they inherit the defects of the univariate AR forecasts in the second period. Second, an examination of the bivariate inflation – activity relation suggests that there have been other changes in that relation, beyond those implied by changes in the univariate inflation process. In particular, the relation between lagged changes in activity and changes in inflation – the lagged effects in the ADL models – are different pre- and post-1984, with the lagged effects being significantly more subdued in the latter period than in the former. One thing which has changed little is that there continue to be significant correlations between changes in inflation and economic activity at business cycle horizons, and the sum of coefficients in backwards-looking Phillips curves specified with levels or gap activity variables tend to be stable and statistically significant using conventional in-sample significance tests. However, these stable longer term relations are too small in magnitude, in an R^2 sense, to produce improvements of simulated real-time activity-based forecasts over the best simulated real-time univariate forecasts over the 1984-2004 period.

The rest of the paper is organized as follows. Section 2 lays out the main forecasting facts. Sections 3 – 5 examine changes in the univariate inflation process, and

Sections 6 and 7 examine changes in bivariate inflation-activity processes. Section 8 concludes.

2. U.S. Inflation Forecasts: Facts, Puzzles, and Hypotheses

This section summarizes the performance of models for forecasting U.S. inflation, using a pseudo out-of-sample forecast comparison methodology. The purpose of this section is to summarize concisely the state of knowledge about changing inflation volatility and forecast performance that has been noted at various points in the recent literature (see Ang, Bekaert, and Wei (2005), Atkeson and Ohanian (2001), Stock and Watson (2002), and Tulip (2005); for complementary results for the UK, see Benati and Mumtaz (2005)). The section begins with a description of the data and the forecasting models, then turns to the results. To keep things simple, in this section we focus on split-sample results, comparing the period 1970:I – 1983:IV to the later period 1984:I – 2004:IV. The sample split date of 1984 coincides with estimates of the onset of the great moderation and is the split date chosen by Atkeson and Ohanian (2001). These split sample results convey the main facts about the changing behavior of inflation forecasts. In later sections, we examine formal evidence for a break and consider methods that allow for continual rather than discrete changes in the inflation process and the forecasting relations.

2.1 Data

The three price indexes used are the GDP deflator (GDPD), the personal consumption expenditure deflator for core items (PCED-core), and the personal consumption expenditure deflator for all items (PCED-all). At different points in this paper, various measures of economic activity are used: the unemployment rate (all, 16+, seasonally adjusted) (u), total housing authorizations as measured by new building permits (permits), GDP, and the capacity utilization rate (manufacturing). For series with revisions, the most recently available vintage (as of May 2005) was used.

All empirical analysis is undertaken with quarterly data. Quarterly values for monthly series were computed by averaging the monthly values for the three months in the quarter; if logarithms are taken, they are logarithms of the average value of the monthly indexes.[1] For the main results, the full sample is from 1959:I through 2004:IV; results that use a different sample are noted explicitly.

Some predictors appear in "gap" form. Gaps are computed as deviation of the univariate activity series (e.g. ln*GDP*) from a lowpass filter with pass band corresponding to periodicities of 60 quarters and higher. Two-sided gaps are computed as deviations from the symmetric two-sided MA(80) approximation to the optimal lowpass filter after padding the endpoints of the series with backcasts and forecasts computed from an estimated AR(4) model. One-sided gaps are computed using the same MA(80) filter replacing future observations with recursively constructed AR(4) forecasts. Two-sided gaps are useful for analyzing historical relationships but are not feasible for forecasting.

2.2 Forecasting Models and Pseudo Out-of-Sample Methodology

We begin by considering two univariate forecasting models and one multivariate forecasting model, implemented using different predictors. Let $\pi_t = 400\ln(P_t/P_{t-1})$, where P_t is the quarterly price index, and let h-period average inflation (at an annual rate) be $\pi_t^h = h^{-1}\sum_{i=0}^{h-1}\pi_{t-i}$. Adopt the notation that subscript $|t$ on a variable denotes the forecast made using data through time t, for example $\pi_{t+h|t}^h$ is the forecast of π_{t+h}^h made using data through t.

AR(AIC). Forecasts are made using a univariate autoregression, specified in terms of the change of inflation with r lags, where r is estimated using the AIC. Multistep forecasts are computed by the direct method, that is, by projecting h-period ahead inflation on r lags. Specifically, the h-step ahead AR(AIC) forecast was computed using the model,

$$\pi_{t+h}^h - \pi_t = \mu^h + \alpha^h(\mathrm{B})\Delta\pi_t + u_t^h, \qquad (1)$$

[1] The analysis was also performed using end-of-quarter aggregation with no important changes in the qualitative conclusions. Many of the coefficient values reported below are sensitive to the method of temporal aggregation (as they should be) but the magnitude and timing of the changes in parameters and the consequent conclusions about forecasting are not.

where μ^h is a constant, α^h(B) is a lag polynomial written in terms of the backshift operator B, u_t^h is the *h*-step ahead error term, and the superscript *h* denotes the quantity for the *h*-step ahead regression. Note that this specification imposes that π_t has a unit root.

AO. Atkeson-Ohanian (2001) (AO) forecasted the average four-quarter rate of inflation as the average rate of inflation over the previous four quarters. They did not forecast at other horizons so there is some ambiguity in specifying the AO forecast at other horizons. Because the AO forecast is essentially a random walk forecast, and a random walk forecast is the same at all horizons, we extend the AO forecast to other horizons without modification. Thus the AO forecast is,

$$\pi_{t+h|t}^h = \pi_t^4 = \frac{1}{4}(\pi_t + ... + \pi_{t-3}) . \quad (2)$$

Backwards-looking Phillips curve (PC). The PC forecasts are computed as direct ADL forecasts, that is, by adding a predictor x_t to (1) to form the autoregressive distributed lag (ADL) specification,

$$\pi_{t+h}^h - \pi_t = \mu^h + \alpha^h(\mathrm{B})\Delta\pi_t + \delta^h(\mathrm{B})x_t + u_t^h , \quad (3)$$

The lag lengths of α^h(B) and δ^h(B) are chosen by AIC. The PC forecast using u_t as a predictor, denoted *PC-u*, is the conventional backwards-looking Phillips curve specified in terms of the level of the unemployment rate with a constant NAIRU, omitting supply shock control variables.[2]

Pseudo out-of-sample forecast methodology. All forecasts were computed using the pseudo out-of-sample forecast methodology, that is, for a forecast made at date *t*, all

[2] A variation on the ADL specification (3) is to allow much longer lags of inflation, but to imposes restrictions that keep the number of free parameters under control, such as in Gordon (1998 and earlier) and Brayton, Roberts, and Williams (1999). Rolling and recursive forecasts constructed using long-lag specifications seemed to produce unstable results and we did not pursue those specifications further.

estimation, lag length selection, etc. was performed using only data available through date *t*.

In this section, we consider recursive forecasts, so that forecasts at date *t* are based on all the data (beginning in 1959:I) through date *t*. The period 1959 – 1970 was used for initial parameter estimation. The forecast period 1970:I – 2004:IV was split into the two periods 1970:I – 1983:IV and 1984:I – 2004:IV.

2.3 Results

The results of the pseudo out-of-sample forecast experiment are summarized in Table 1, where the different panel of the table report results for the three inflation series. The first row in each panel reports the root mean square forecast error (RMSFE) of the benchmark AR(AIC) forecast in percentage points at an annual rate, at the indicated forecast horizon *h*. The remaining rows report the MSFE of the row forecast, relative to the AR(AIC) (so the relative MSFE of the AR(AIC) forecast is 1.00); an entry less than one indicates that the candidate forecast has a lower MSFE than the AR(AIC) benchmark.

Four findings from Table 1 stand out.

1. *The RMSFE of inflation forecasts has declined, and in this sense inflation has become easier to forecast*. The magnitude of this reduction bears emphasis. Whatever its other merits or demerits, the AR(AIC) forecast is simple to produce and has been a staple of economic forecasters for decades, and a forecaster using this method consistently from 1984 to 2004 would have forecast errors of only 0.56 percentage points for annual core PCE inflation, down from 1.66 percentage points over the 1970-1983 period, a reduction of two-thirds. As the final column of Table 1 demonstrates, even those forecasting models with performance that deteriorates in the second period, relative to the first, still provide considerable reductions in forecast uncertainty of at least one-half.

2. *The relative performance of the Phillips curve forecasts deteriorated substantially from the first period to the second.* For example, during the 1970-1983 period at the four-quarter horizon, the PC-*u* forecast of PCE-core outperformed the AR(AIC) benchmark (relative MSFE = .94), but during the 1984-2004 period it performed worse than the AR(AIC) benchmark (relative MSFE = 1.32). The change in relative performance is even larger at $h = 8$, but there are fewer nonoverlapping observations at this horizon so the $h = 8$ relative MSFEs have considerable sampling uncertainty. This deterioration of Phillips curve forecasts is found not just for the activity predictors examined in the table (unemployment and building permits), but for other activity predictors as well (including the capacity utilization rate and coincident indexes of real activity).

3. *The poor performance of the PC forecasts is not simply a consequence of failing to allow for a time-varying NAIRU.* The PC-Δu, PC- $u-\bar{u}^{1-Sided}$, and PC-Permits specifications all allow for a slowly time-varying NAIRU: the PC-Δu specification, by omitting the levels term; the PC- $u-\bar{u}^{1-Sided}$, by using a gap based on one-sided detrending, and the PC-Permits specification by using a stationary predictor. In some cases these outperform the PC-*u* specification, in other cases they do worse, but in any event none produces multiquarter forecasts that are competitive with either the AR(AIC) or AO forecasts.

4. *The AO forecast substantially improves upon the AR(AIC) and Phillips curve forecasts at the four- and eight-quarter horizons in the 1984-2004 period, but not at shorter horizons and not in the first period.* The shift in relative performance is found for all three price indexes and is dramatic. For example, at the $h = 4$ horizon (the only horizon reported by Atkeson and Ohanian (2001)), in the first period the MSFE of the AO forecast, relative to the PC-*u* forecast, is 1.13/.99 = 1.14, whereas in the second period this relative MSFE is .74/1.26 = 0.59.

Evidently, there have been major changes in the univariate inflation process and in the bivariate process of inflation and its activity-based predictors, however these results do not indicate what those changes were or when they occurred.

3. Changes in the Univariate Inflation Process

Our study of changes in inflation forecasting models begins with the univariate inflation process. This section continues to consider split-sample results with a 1984 break date; timing of changes is considered in Section 4.

3.1 Autocorrelations, Spectra, and Measures of Persistence

There have been substantial changes in the autocorrelations and spectra of inflation. Some measures of persistence of the inflation process have changed, while others have not.

Autocorrelations and spectra. Table 2 presents the first eight autocorrelations of $\Delta\pi_t$ for the three inflation measures in both periods, along with its standard deviation.

Two aspects of these autocorrelations are noteworthy. First, for PCE-core and PCE-all, the only statistically nonzero autocorrelation (at the 5% level) is the first; this is true in both periods. For the GDP deflator, in the first period the only autocorrelation that is nonzero at the 10% level is the first (the *t*-statistic is 1.89); in the second period, the first autocorrelation is statistically significant at the 5% level, but so is the fourth, perhaps reflecting some seasonality. For all series, the first autocorrelation is negative. The second noteworthy aspect is that the first autocorrelation is much larger in absolute magnitude (more negative) in the second period than the first, for all three series.

These results suggest that PCE-core and PCE-all inflation might be well described by an integrated moving average model with a single MA term, that is, by an IMA(1,1) model in both periods, with a MA coefficient that is larger in absolute magnitude in the second period than in the first. The IMA(1,1) model also is a plausible candidate for the GDP deflator in the first period.

The estimated spectrum of the change in PCE-all inflation is plotted in Figure 1. Two estimates are reported: a nonparametric estimator (smoothed periodogram) and a parametric IMA(1,1) estimator. The parametric estimate looks like a smoothed version of the nonparametric estimate, suggesting that the IMA(1,1) model fits the data reasonably well. Relative to the first period, the spectrum in the second period is lower in magnitude – this reflects the reduction in volatility between the two periods – and has more power at high than at low frequencies. Closer inspection reveals that the shape of the spectrum has changed, as well as its level, with the second period having relatively more power at higher frequencies than in the first. This is consistent with the more negative first autocorrelation in the second period than in the first.

Persistence. One measure of persistence is the magnitude of the largest autoregressive root of the levels process, in this case, the largest autoregressive root of inflation. As shown in Table 3, by this measure persistence did not change substantially between the two periods, either in a qualitative or quantitative sense. The confidence intervals are remarkably stable across series and time periods, being approximately (.85, 1.03); all the confidence intervals include a unit root, and all exclude roots less than .83. By this measure, the persistence of inflation has not changed between the two periods. These split-sample confidence intervals are consistent with the results of the thorough analysis in Pivetta and Reis (2004), who report rolling and recursive estimates of the largest autoregressive root (and the sum of the AR coefficients) that are large and stably near one.

Although the largest AR root appears to be stably large, inflation persistence can nevertheless be viewed as having fallen: the fraction of the variance of $\Delta\pi_t$ explained by the persistent shocks and the fraction of the mass of the spectrum of $\Delta\pi_t$ near frequency zero are both considerably greater in the first period than in the second.

3.2 Trend-Cycle Decomposition

The IMA(1,1) and unobserved components models. The apparent unit root in π_t and the negative first order autocorrelations, and generally small higher order autocorrelations, of $\Delta\pi_t$ suggest that the inflation process can be described by the IMA(1,1) process,

$$\Delta\pi_t = (1 - \theta B)a_t, \tag{4}$$

where a_t is serially uncorrelated with mean zero and variance σ_a^2. Equivalently, π_t can be represented as an unobserved components (UC) model with a stochastic trend τ_t and a serially uncorrelated disturbance η_t:

$$\pi_t = \tau_t + \eta_t, \quad \eta_t \text{ serially uncorrelated } (0, \sigma_\eta^2) \tag{5}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t, \quad \varepsilon_t \text{ serially uncorrelated } (0, \sigma_\varepsilon^2), \tag{6}$$

where $\text{cov}(\eta_t, \varepsilon_t) = 0$.

Results. Table 4 presents estimates of the IMA(1,1) parameters, the implied UC parameters, and statistics testing the IMA(1,1) model against more general ARIMA models, for the pre- and post-1984 periods. Consistent with the changes in the autocorrelations, the MA parameter is considerably larger in the second period than in the first; consistent with the decline in the variance of inflation and the decrease in the power of the spectrum at all frequencies, the MA innovation has a smaller variance in the second period than in the first. This is true for all series.

The final three blocks of table 4 report various statistics assessing the fit and stability of the IMA(1,1) model. As can be seen in the third block in the table, Wald tests of the null that the process is an IMA(1,1), against the alternative that it is a higher order process, fail to reject in all cases, with a single anomalous result favoring an IMA(1,4) specification for PCE-all in the first period.

Although the unit root confidence intervals in table 3 include one, the confidence intervals include values of the largest AR root that are relatively small, less than 0.9, so one might ask for additional evidence on the magnitude of the AR root in the context of a model with a moving average term. To this end, table 4 reports ARIMA(1,0,1) models that do not impose a unit root in inflation. The point estimates are strikingly close to one, the smallest being .986. Because the distribution theory for the estimator of α is nonstandard when its true value is close to one, we rely on table 3 for formal inference about this root.

The fifth block of the table reports tests of the hypothesis of parameter stability in the UC model. The hypothesis that the permanent innovation has the same variance in the two periods is strongly rejected, but the hypothesis that the transitory variance is the same is not. These tests are Chow tests which treat the 1984 break date as exogenously specified, which is inappropriate in that a large body of evidence about changes in the U.S. macroeconomy informed our choice of a 1984 break. To address this concern, the final line reports the Quandt likelihood ratio (QLR) statistic, which is the maximum likelihood ratio test for a break over all possible break dates in the inner 70% of the full sample (1959 – 2004); for all series, the QLR test rejects at the 1% significance level, providing formal evidence of instability in the parameters of the UC model. The tests on individual parameters suggest that this instability appears in the permanent innovation variance but not the transitory innovation variance.

One measure of how important the permanent shocks are is to decompose the four-quarter ahead forecasts into three sources: errors in estimation of the current trend, that is, signal extraction (filtering) errors, forecast errors arising from currently unknown permanent disturbances over the next four quarters, and forecast errors arising from currently unknown transitory disturbances over the next four quarters. This decomposition is given in the final block of Table 4. For all three inflation series, in the first period future trend disturbances are by far the largest source of four-quarter forecast errors, followed by filtering errors. The magnitude of both sources of error reduces dramatically for all three series from the first period to the second: the forecast error attributed to the trend disturbance falls by at least 90% for all three series, and the forecast error variance arising from filtering error falls by between one-third and two-thirds. Like the contribution of the trend shock itself, the decline in the contribution of the filtering error is a consequence of the decline in the volatility of the trend shock because the trend is less variable and therefore is estimated more precisely by the UC filter. The contribution of the transitory shocks remains small and is approximately unchanged between the two periods.

Historical precedents. The IMA(1,1) representation for inflation is not new. As was mentioned in the introduction, Nelson and Schwert (1977) selected an IMA(1,1) model for monthly U.S. CPI inflation (identified, in the Box-Jenkins (1970) sense, by inspecting the autocorrelogram of π_t and $\Delta\pi_t$). Using a sample period of 1953m2 – 1971m7, they estimated an MA coefficient of .892 (their equation (4)). This monthly IMA(1,1) model temporally aggregates to the quarterly IMA(1,2) model[3],

$$\Delta\pi_t = (1 - .487\mathrm{B} - .158\mathrm{B}^2)a_t. \qquad \text{(NS77)} \qquad (7)$$

Schwert (1987) reports IMA(1,1) models for two monthly price indexes (the CPI and the PPI) and one quarterly index, the GNP deflator. Using data from 1947:I – 1985:IV, Schwert (1987, Table 6) estimated the MA coefficient for the GNP deflator to be .665.

A third historical reference is Barsky (1987), who uses Box-Jenkins identification methods to conclude that quarterly CPI inflation (third month of quarter aggregation) is well described by an IMA(1,1) model, with a MA coefficient that he estimates to be .46 over the 1960-1979 period (Barsky (1987, Table 2)).

We return to these historical estimates below.

3.4 Explaining the AO Results

All the univariate models considered so far, including the AO model, produce multistep forecasts that are linear in π_t, π_{t-1},…, so one way to compare these models is to compare their forecast functions, that is, their weights on π_t, π_{t-1},…. Figure 2 plots the

[3] Let p_t^Q be the quarterly log price series, t indexed in months, and let P_t^m be the monthly log price index. Suppose that the GDP deflator is best thought of as representing average prices during the quarter (not end-of-quarter prices). Accordingly, for comparability to the GDP deflator, we compute the quarterly PCE log price index as $p_t^Q = \ln((P_t^m + P_{t-1}^m + P_{t-2}^m)/3)$. Then quarterly inflation is $(1 - \mathrm{B}^3)\, p_t^Q$ and the first quarterly difference of inflation is $(1 - \mathrm{B}^3)^2 p_t^Q$. If $\Delta\ln(P_t^m)$ follows the monthly IMA(1,1) process $\Delta^2\ln(P_t^m) = (1 - \theta^m\mathrm{B})a_t$, then the monthly latent quarterly inflation series follows $(1 - \mathrm{B}^3)^2 p_t^Q = (1 + \mathrm{B} + \mathrm{B}^2)^3(1 - \theta^m\mathrm{B})a_t$, which, when sampled every third month, corresponds to an IMA(1,2) at the quarterly frequency.

forecast functions for four quarter ahead forecasts computed using the AO model, an IMA(1,1) model with $\theta = .25$, and an IMA(1,1) model with $\theta = .65$. The value $\theta = .25$ closely approximates the value of θ estimated using the 1970-1983 sample for all three series, reported in Table 4, and the value $\theta = .65$ closely approximates the estimates obtained using the 1984-2004 sample. The AO forecast function weights the most recent four quarters of inflation evenly, whereas the IMA(1,1) forecast functions are geometrically declining. The AO and $\theta = .25$ forecast functions are quite different, and the resulting forecasts typically would be quite different. In contrast, the AO and $\theta = .65$ forecast function provides a closer approximation to the $\theta = .65$ forecast function, and one might expect the AO and $\theta = .65$ forecasts to be fairly close much of the time.

The changing coefficients in the IMA(1,1)/UC representation thus provide a concise arithmetic explanation for the performance of the AO forecasts evident in Table 1. Over the 1970-1983 period, during which the MA coefficient is small, the AO model would be expected to work poorly. During the later period, during which the MA coefficient is large, at the four-quarter horizon the AO model provides an approximation to the IMA(1,1) forecast and would be expected to work well. This approximation is also good at the eight-quarter horizon, but not at short horizons, so the AO model would be expected to work well at longer seasonal horizons, but not short horizons in the second period. This pattern matches that in Table 1.

4. An Unobserved Components – Stochastic Volatility Model of Inflation

The tests for parameter instability reported in Table 4 indicate that there have been statistically significant and economically large changes in the univariate inflation process. This raises the question of when those changes occurred and whether they are associated with continual parameter drift or discrete regime shifts.

In this section, we examine the evolution of the inflation process by generalizing the unobserved components model of the previous section to allow the variances of the permanent and transitory disturbances to evolve randomly over time. Specifically, we model the logarithms of the variances of η_t and ε_t as evolving as independent random walks. The result is the unobserved components model with stochastic volatility (UC-SV),

$$\pi_t = \tau_t + \eta_t, \quad \text{where } \eta_t = \sigma_{\eta,t}\zeta_{\eta,t} \tag{8}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t, \quad \text{where } \varepsilon_t = \sigma_{\varepsilon,t}\zeta_{\varepsilon,t} \tag{9}$$

$$\ln\sigma^2_{\eta,t} = \ln\sigma^2_{\eta,t-1} + \psi_{\eta,t} \tag{10}$$

$$\ln\sigma^2_{\varepsilon,t} = \ln\sigma^2_{\varepsilon,t-1} + \psi_{\varepsilon,t} \tag{11}$$

where $\zeta_t = (\zeta_{\eta,t},\ \zeta_{\varepsilon,t})$ is i.i.d. N(0, I_2), $\psi_t = (\psi_{\eta,t},\ \psi_{\varepsilon,t})$ is i.i.d. N(0, γI_2), and ζ_t and ψ_t are independently distributed, and γ is a scalar parameter. Note that this model has only one parameter, γ, which controls the smoothness of the stochastic volatility process; γ can either be estimated or chosen *a-priori*.

Results. Figures 3 plots the smoothed estimates of $\sigma_{\eta,t}$ and $\sigma_{\varepsilon,t}$ from the UC-SV model, computed by MCMC using a vague prior for the initial condition and $\gamma = 0.2$, for the GDP deflator, using data from 1953:I – 2004:IV (the longer sample is used here to obtain estimates for the 1950s, and to facilitate comparisons with the Nelson-Schwert (1977) estimate, which was based on data starting in 1953).

The estimates in Figure 3 show substantial time variation in the standard deviation of the permanent component: the 1970s through 1983 was a period of high volatility, 1953 through the late 1960s or early 1970s and 1984-1990 were periods of moderate volatility of the permanent innovation, and the 1990s through 2004 have been period of low volatility of the permanent innovation. In contrast, there is little variation in the estimates of the variance of the transitory innovation. In consequence, the time variation of the implied coefficient in the implied instantaneous IMA(1,1) representation tracks inversely the smoothed estimates of $\sigma_{\varepsilon,t}$, being moderate (around .4) in the 1950s through late 1960s, small (less than .25) during the 1970s through 1983, higher in the late 1980s, and increasing further in the 1990s to a current estimate of approximately .85.

Figure 4 presents smoothed estimates of the instantaneous standard deviations for PCE-core and PCE-all; the results are very similar to those in figure 3.

A UC-SV model with heavy-tailed volatility innovations. The UC-SV model specifies the log variances as following a Gaussian random walk. This imparts smoothness to the stochastic volatility, relative to a stochastic volatility process with heavier tails. If a regime shift model is a better description of the changes in volatility than a model, like the UC-SV model, with smooth parameter variation, then the UC-SV model implemented above might miss the rapid changes in volatility associated with a shift in regimes.

To investigate this possibility, we modified the UC-SV so that the disturbances $\zeta_{\eta,t}$ and $\zeta_{\varepsilon,t}$ were drawn from a mixture of normal distributions, N(0,.1I) with probability .95 and N(0,.5I) with probability .05; the second part of this mixture creates a heavy-tailed mixture distribution with occasional large jumps. As it happens, the smoothed estimates of $\sigma_{\eta,t}$ and $\sigma_{\varepsilon,t}$ from the mixture-of-normals UC-SV model are qualitatively and quantitatively close to those from the normal-error UC-SV model. To conserve space, the mixture-of-normal UC-SV results are not presented, and for simplicity the only UC-SV model considered below is the normal-error model (8) – (11).

Relation to estimates in the literature. The series analyzed by Nelson and Schwert (1977), Schwert (1987), and Barsky (1987) differ from those analyzed here, and the Nelson-Schwert and Schwert estimates for CPI were computed using monthly data. Despite these differences, their estimated IMA(1,1) parameters are consistent with figures 3 and 4. The published estimates of the quarterly MA coefficient are .46 for 1960-1979 (Barsky (1987)) and .665 for 1949-1985 (Schwert (1987)). Comparing these estimates to averages in figure 3(c) over the corresponding time period indicates general agreement between the instantaneous MA coefficient estimated in figure 3(c) and the earlier estimates (the high estimate in Schwert (1987) seems to be driven in part by the pre-1953 data).

5. Pseudo Out-of-Sample Univariate Forecasts

This section examines whether the UC model with time-varying coefficients could have produced useful real-time univariate forecasts by extending the pseudo out-of-sample forecasting results, reported in Section 2, to include additional UC models.

5.1 Forecasting Models

The comparison considers two models already reported in Table 1, the recursive AR(AIC) benchmark and the AO model. The comparison also considers the following additional models:

Recursive IMA(1,1) and AR(4). The models are estimated using an expanding sample starting in 1959.

Rolling AR(AIC), IMA(1,1), and AR(4). The models are estimated using a data window of forty quarters, concluding in the quarter of the forecast.

Nelson-Schwert (1977) (NS77). The NS77 model is the quarterly IMA(1,2) model (7) implied by temporal aggregation of Nelson and Schwert's (1977) monthly IMA(1,1) model.

UC-SV, γ=.2. Forecasts are computed using the UC-SV model with $\gamma = .2$. The UC-SV model is applied to data from 1959 through the forecast quarter to obtain filtered estimates of the current values of $\sigma_{\varepsilon,t}$ and $\sigma_{\eta,t}$, which in turn are used to construct the forecasts.

Fixed coefficient IMA(1,1). These are IMA(1,1) models with coefficients of .25 and .65. The coefficient of .25 approximately corresponds to the value in Table 2 for the period 1970 – 1983, and the coefficient of .65 approximately corresponds to the value for 1984 – 2004.

Multiperiod forecasts based on the IMA(1,1) and UC models are iterated, and multiperiod AR forecasts use the direct method (1).

The recursive and rolling models produce pseudo out-of-sample forecasts. The NS77 model produces a true out of sample forecast, since the coefficients were estimated using data through 1971. The UC-SV model with $\gamma = .2$ and the fixed-coefficient IMA(1,1) models are not pseudo out-of-sample models because their parameters (γ in the first instance, θ and σ_a in the second) were estimated (or, in the case of γ, calibrated) using the full data set, so in particular their parameters were not estimated by recursive or rolling methods.

5.2 Results

Table 5 summarizes the forecasting performance of the various models over the 1970-1983 and 1984-2004 periods (entries are MSFEs, relative to the recursive AR(AIC), which is also the benchmark in Table 1). Figures 5-7 provides additional detail about the forecasting performance at different points in time by presenting a two-sided smoothed estimate of the relative MSFE (relative to the recursive AR(AIC) forecast), with exponential smoothing and a discount factor of .95 (end points are handled by simple truncation).

Inspection of table 5 and figures 5-7 suggests four findings.

First, among the fixed-parameter models, the $\theta = .25$ model performs well in the first period and poorly in the second, whereas the $\theta = .65$ model performs poorly in the first period and well in the second. This is consistent with the choice of these two parameter values as being approximately the MLEs of θ in the two periods. More notably, the UC-SV(.2) model evidently adapts well to the shifting parameter values and produces forecasts that rival those of the $\theta = .25$ model in the first period, and the $\theta = .65$ model in the second. Accordingly we will treat the UC-SV(.2) model can be thought of as providing an approximate bound on the forecasting performance of the pseudo out-of-sample forecasts.

Second, among the pseudo out-of-sample forecasts, the rolling IMA(1,1) model has the best performance, or nearly so, for all series and at all horizons. For core PCE, the rolling IMA(1,1) model has the best average performance in both the 1970-1983 and 1984-2004 periods in both horizons. For the GDP deflator and PCE-all, the IMA(1,1) forecast is best in the first period and is only slightly worse than the AO forecast at the four- and eight-quarter horizons in the second period. Closer inspection of the forecast errors indicates that the primary source of the improvement of the AO model over the rolling IMA(1,1) model during the second period occurs during the late 1980s, a period of sharp change in the MA coefficient during which the rolling forecast took time to adapt.

Third, the autoregressive forecasts are notably worse than the IMA/UC forecasts. This is consistent with inflation being well described by an IMA with a changing coefficient that is large in the second period. If the IMA model is a good approximation, then the AR(AIC) models will suffer from approximation error and, if the MA coefficient is large, will tend to select large lag lengths and therefore suffer from additional parameter estimation error, resulting in forecast deterioration. Eliminating the AIC lag selection by using an AR(4) fails to improve performance; this is also true if AIC lag selection is replaced by BIC (unreported results).

Fourth, the NS77 forecast is truly an out-of-sample forecast, and its performance is remarkable. During the 1970-83 period, its performance was comparable to the AR(AIC) benchmark at short horizons (for the GDP deflator, somewhat better, for PCE-core and PCE-all, somewhat worse). At the two-year horizon in the first period, the NS77 model works very well for all series. The most striking feature of the NS77 results is that the NS77 model has extremely good performance in the second period, both at short and long horizons, in fact beating the other models – including the AO model – for seven of the twelve series/horizon combinations considered.

5.3 Summary

The results in Sections 3 – 5 provide a simple picture of the evolution of the inflation process. The permanent component of inflation increased in magnitude during the 1970s through 1983, then fell significantly, both in a statistical and economic sense. Although the stochastic trend component of inflation diminished in importance, it remains nonzero (confidence intervals for the largest AR root include one). Moreover, because the variance of the permanent component (but not the transitory component) fell dramatically, the variance of $\Delta\pi_t$ has fallen. Because the smaller trend variance corresponds to a larger MA coefficient, the declining importance of the stochastic trend in inflation explains both the good performance of the AO forecast in the second sample at long horizons, and its poor performance at short horizons and in the first sample. An important piece of evidence supporting this interpretation is that the rolling IMA(1,1)

(UC) model produces forecasts that have the lowest MSFE, or nearly so, at all horizons for all three series among the recursive and rolling univariate models, including the AO model.

6. Changes in Multivariate Models

We now turn to an examination of the changing relationship between inflation and various measures of economic activity, first at business cycle frequencies, then in autoregressive distributed lag (ADL) models (backwards-looking Phillips curves). The analysis in this section focuses in-sample statistics; we return to forecasting in Section 7.

Tables 6 and 7 summarize different aspects of the bivariate relationships between $\Delta\pi_t$ and various activity measures (x_t) in the two periods. In both tables, the statistics are computed for two periods, 1960-1983 and 1984-2004, using full-subsample estimates within each subsample (no recursive or rolling estimation). The statistics reported in table 6, which were computed from bivariate VAR(4)s for $\Delta\pi_t$ and x_t, are the MSEs of the four-quarter ahead (iterated) bivariate prediction, relative to the ARIMA(8,1,4) univariate model for π_t implied by the VAR(4), and second the coherence between $\Delta\pi_t$ and x_t, averaged over the business cycle frequencies of 6 to 24 quarters. Table 7 reports the ADL(4,4) regressions of the form (3), which are the direct counterpart of the VARs reported in Table 6. The statistics presented in table 7 are the implied (in-sample) four quarter ahead MSE of the ADL forecast, relative to the four-quarter ahead direct AR(4) forecast, the sum of regression coefficients on the gap variable x_t (this is omitted if the specification is in terms of Δx_t), the four-step Granger-causality F-statistic, and various statistics testing for constancy of subsets and all of the coefficients across the two periods.

The results in tables 6 and 7 suggest that some aspects of the inflation-output relation are stable between the two periods, but many aspects are not.

Many things have changed… The relative MSEs in the two tables all show a decline in the marginal predictive content of the activity variables from the first period to the second: typical relative MSEs in the first period are .7, while typical relative MSEs in the second are .9. These declines are consistent with the deterioration of ADL activity-based forecasts between the two periods documented in Section 2. Note that in both tables 1 and 2, the relative MSEs have the interpretation of $1–R^2$ for that horizon: they are in-sample, are not adjusted for degrees of freedom or estimation uncertainty, and the relative MSEs cannot exceed one by construction. The combined effects of parameter instability and estimation error means that the in-sample relative MSEs of .9 in the second period are consistent with relative MSFEs for pseudo out-of-sample forecasts exceeding one, as they do in table 1.

The bivariate $\Delta\pi_t$, Δu_t relationship is of particular interest, and one estimate of the changes in this relationship is provided by the impulse response functions in each of the two periods implied by a VAR(4). These are shown in figure 8, where the impulse responses correspond to the "inflation first" Cholesky ordering. The impulse responses with respect to the inflation shock change somewhat between the two periods with less persistence in the second period. The impulse responses with respect to the output shock exhibit considerable changes. In particular, the response for the first several quarters to an output shock is similar in the two periods, but the output shock has more than twice the effect on inflation in the first period than in the second over the four- to twelve-quarter horizon.

Table 8 reports variance decompositions for the four-quarter ahead inflation forecasts constructed from regressions of the period-specific univariate IMA(1,1) forecast errors onto four lags of the activity predictors and a constant. The univariate forecast errors are first decomposed into the explained and unexplained components within each period, and the change in the explained component is further decomposed into the change associated with the sample variance of the regressors and the change associated with the regression coefficients. Although the variance of the dependent variable has fallen substantially (reduced volatility of four-quarter inflation and smaller univariate forecast errors), the amount of that variability explained by the activity predictors has fallen even more, producing a decline in the R^2. Some of the reduction in the explained sum of squares is accounted for by the reduction in the volatility of the activity measure (the great moderation effect), but the variance of inflation fell by more than the variance of the activity measures so the rest of the decline is a consequence of changes in the coefficients associated with the predictors, in particular (drawing upon the Chow tests reported in table 7), changes in the coefficients on lagged Δx_t.

But some have not… The relationship between annual or longer changes in inflation and the level of the output gap – the backwards-looking Phillips curve relationship – exhibits some stability. Despite the clear change in the VAR coefficients, the implied coherence between the activity variable and changes of inflation at business cycle frequencies is large and relative stable across periods, with most of the estimates in Table 6 falling in the range 0.5 – 0.6. This stable coherence is consistent with the positive and stable association found at business cycle frequencies by Harvey, Trimbur, and van Dijk (2005, figures 16 and 17), who used a bivariate unobserved components (trend-cycle) model with different, but possibly correlated, real and nominal cyclical components.

Similarly, the sum of coefficients on the gap variables in the ADL specifications in table 7 are fairly constant, at least for the GDP deflator and PCE-core. In none of the six cases in table 7 (unemployment gap and output gap, three price indexes) is the hypothesis of constancy of the sum of coefficients rejected at the 10% level; this despite the rejection of the constancy of all the ADL coefficients at the 5% level (five at the 0.1% level), in all six specifications. The ADL Chow test statistics suggest that the coefficients on lagged Δx have changed, but the coefficients on the levels (in gaps) have not.

Additional evidence of some stability in the relationship between four-quarter inflation and gap measures is presented in figure 9, which plots the four-quarter ahead prediction error from the UC-SV model ($\gamma = .2$) and the two-sided unemployment gap; the regression lines corresponding to the scatterplots in figure 9 (and also for the output gap) are presented in table 9. For all series, the slope of this bivariate regression is less in the second period than in the first, and the $\bar{R}^2$ falls by roughly one-half between the two periods. This said, the slope remains statistically significant (and negative for the unemployment gap, positive for the output gap) for the GDP deflator and core PCE.

7. Pseudo Out-of-Sample Multivariate Forecasts

The univariate analysis indicated that, during the second sample period, inflation dynamics were approximated less well by an autoregression than by an IMA model with a large MA coefficient. Because the ADL model extends the AR model to additional predictors, this raises the possibility that a different framework, built upon the IMA model instead of the ADL model, might provide a more suitable foundation for including an activity variable. Accordingly, in this section we consider the pseudo out-of-sample forecasting performance of an alternative direct forecast method, in which an activity variable is used to predict the dependent variable is the four-quarter ahead forecast error from the UC-SV model. Specifically, in this section we consider forecasting regressions of the form,

$$e^4_{t+4|t} = \mu + \alpha\Delta\pi_t + \delta(\mathrm{B})x_t + u^4_{t+4}, \qquad (12)$$

where $e^4_{t+4|t}$ is the four-quarter ahead forecast error from the UC-SV model with $\gamma = .2$.

Pseudo out-of-sample forecasting results for rolling estimates of equation (12) are summarized in Table 10, for specifications in which $\Delta\pi_t$ is excluded (odd numeric columns) from (12) and for specifications in which $\Delta\pi_t$ is included, for both a single lag of x_t and for BIC choice of lags. The numeric entries are MSFEs, relative to the UC-SV forecast (which corresponds to all coefficients in (12) being zero), so an entry less than one indicates a pseudo out-of-sample improvement using x_t (and, in the odd columns, $\Delta\pi_t$) in (12), relative to simply using the UC-SV forecast.

The results in table 10 are negative. In the few combinations of activity variables and inflation indexes in which the pseudo out-of-sample relative MSFE is slightly less than one, the improvement is not robust to changing the inflation series. In some cases, the rolling forecasts perform far worse than the UC-SV benchmark, and far worse than would be expected by sampling error alone were the true distributed lag $\delta(\mathrm{B})$ in (12) stably zero, pointing to instability of the activity coefficients. This said, comparing forecasts for the same predictor and inflation series in tables 1 and 10 reveals that using the imposing the UC-SV dynamics improves the performance of the activity-based forecast. For example, the ADL four-quarter ahead forecast of core PCE, denoted by PC-Δu in table 1, has a MSFE, relative to the UC-SV model of 1.59[4], whereas for specification (12) the relative MSFE is 1.06; when the one-sided gap is used as the activity variable, the relative MSFE for the four-quarter ahead ADL forecast is 2.02, whereas it is 1.12 for the forecast based on (12).

One notable feature of the results in table 10 is that using one-sided gaps or levels variables does not improve forecasting performance relative to dropping the gaps and simply using first differences (compare the *u*-gap to the Δu results, and the *GDP* gap to the $\Delta \ln GDP$ results), although using gaps does seem to mitigate the largest forecast errors. Thus the fairly stable in-sample correlations in table 9 do not appear to be large enough (or, possibly, stable enough) to help in this pseudo out-of-sample forecasting comparison.

The negative results in table 10 do not show that it is impossible to improve upon the UC-SV or rolling IMA(1,1) forecasts, just that it is not easy to do so within the context of linear models with a single additional activity predictor. In fact, there is evidence that real-time forecasts have provided improvements upon the best univariate models. Kohn (2005) reports large reductions in true (not pseudo) mean squared errors of real-time Fed staff forecasts of quarterly CPI inflation, relative to the AO model, over 1984-2000, and Ang, Bekaert, and Wei (2005) find that true out-of-sample survey forecasts (median Michigan or Livingston forecasts) outperform a large number of pseudo out-of-sample univariate and multivariate time series competitors. Both the Board staff forecast and the survey forecasts are combination forecasts, pooled over judgmental and model-based forecasts, and both presumably incorporate considerably more information than are present in the simple activity-based forecasts examined here. Perhaps further attempts to develop competitive time series forecasts could profit from pursuing systematically those features that have proven successful in these survey forecasts.

[4] 1.59 = 1.45 (from table 1, which is relative to AR(AIC)) ÷ .91 (UC-SV relative to AR(AIC), from table 5).

8. Discussion and Conclusions

This paper has focused on providing a mechanical interpretation of the changes in the inflation process, and in bivariate inflation-activity processes, that have led to the breakdown of inflation forecasts made using autoregressions and ADL-based backward-looking Phillips curves. We have tried to provide a more nuanced and specific understanding of these changes than is provided by a headline conclusion about the disappearance of the (backward-looking) Phillips curve, a headline that is accurate in parts, inaccurate in parts, and masks a number of interesting developments. Our proposed explanation has three parts: first, that the variance of permanent disturbances to inflation has changed considerably over time (and today is at historically low levels); second, that past changes in activity variables have substantially less predictive content for annual inflation now (and over the past twenty years) than they did during the 1970s; and third, the centerpiece of the backwards-looking Phillips curve, the relationship between changes in inflation and activity gaps over the business cycle, seems to have been fairly stable, but this correlation is so small in an R^2 sense that it is difficult to exploit it to improve out-of-sample forecasts (at least, the attempts to do so in this paper were unsuccessful).[5]

One thing this paper has not done is to attempt to link these changes in time series properties to more fundamental changes in the economy. The obvious explanation is that these changes stem from changes in the conduct of monetary policy in the post-1984 era, moving from a reactive to a forward-looking stance (see for example the recent

[5] Our analysis has focused on the backwards-looking Phillips curve because the backwards-looking specification is the one relevant for forecasting, the topic of this paper. The changes documented here have no direct implications for forward-looking New Keynesian Phillips curve specifications, which may or may not have been unstable. (See Li (2005) for a discussion of various forward looking specifications.) The results here *do* have implications for uncovering instability in NKPC specifications, however, since the multivariate specifications considered here are close to (or in some cases the same as) the first stage in instrumental variables estimations of the NKPC parameters. There are already concerns about those parameters being weakly identified (e.g. Mavroedis (2004)). The results here suggest that they are particularly weakly identified in the post-1984 sample, so that quantifying changes in the coefficients is both challenging (because of weak instruments) and may be uninformative (because stability tests my have low power).

discussion in Estrella (2005), who explains the post-80s failure of the term structure to have predictive content for inflation in terms of changes in Fed policy). But obvious explanations are not always the right ones, and there are other possibilities. To a considerable extent, these other possibilities are similar to the ones raised in the context of the discussion of the great moderation, including changes in the structure of the real economy, the deepening of financial markets, and possible changes in the nature of the structural shocks hitting the economy. We do not attempt to sort through these explanations here, but simply raise them to point out that the question of deeper causes for these changes merits further discussion.

References

Ang, A., G. Bekaert, and M. Wei (2005), "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?" manuscript, Columbia University.

Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica* 61:821-856.

Atkeson, A. and L.E. Ohanian (2001), "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review* 25(1):2-11.

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

Benati, L. and H. Mumtaz (2005) "The 'Great Stability' in the U.K.: Good Policy or Good Luck?" manuscript, Research Department, Bank of England.

Box, G.E.P. and G.M. Jenkins (1970), *Time Series Analysis*. San Francisco.

Brayton, F., J.M. Roberts, and J.C. Williams (1999), "What Happened to the Phillips Curve?," FEDS Working Paper 1999-49.

Estrella, A. (2005), "Why Does the Yield Curve Predict Output and Inflation," *The Economic Journal* 115:722-744.

Gordon, R.J (1998), "Foundations of the Goldilocks Economy: Supply Shocks and the Time-Varying NAIRU," *Brookings Papers on Economic Activity* 1998(2):197-333.

Harvey, A.C., T.M. Trimbur, and H.K. van Dijk (2005), "Trends and Cycles in Economic Time Series: A Bayesian Approach," manuscript, Cambridge University, UK.

Kohn, D.L. (2005), "Modeling Inflation: A Policymaker's Perspective," speech to the International Research Forum on Monetary Policy, Frankfurt, Germany, May 20, 2005, available at http://www.federalreserve.gov/boarddocs/speeches/2005/20050520/default.htm.

Li, Hong (2005), "Testing Alternative Models of Price Adjustment," mimeo, Brandeis University.

Mavroedis, S. (2004), "Weak Identification of Forward-Looking Models in Monetary Models," *Oxford Bulletin of Economics and Statistics* 66, Supplement, 609-635.

Nelson, C.R. and G.W. 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* 67:478-486.

Pivetta, F. and R. Reis (2004), "The Persistence of Inflation in the United States," manuscript, Princeton University.

Stock, J.H. (1991), "Confidence Intervals for the Largest Autoregressive Root in U.S. Economic Time Series," *Journal of Monetary Economics*, 28:435-460.

Stock, J.H., and M.W. Watson (1999), "Forecasting inflation," *Journal of Monetary Economics* 44:293-335.

Stock, J.H. and M.W. Watson (2002), "Has the Business Cycle Changed and Why?" *NBER Macroeconomics Annual* 2002, 159–218.

Stock, J.H., and M.W. Watson (2003), "Forecasting output and inflation: The role of asset prices," *Journal of Economic Literature* 41:788-829.

Tulip, P. (2005), "Has Output Become More Predictable? Changes in Greenbook Forecast Accuracy," FEDS no. 2005-31.

Table 1
Pseudo Out-of-Sample Forecast Results

A. GDP Deflator

	1970:I – 1983:IV				1984:I – 2004:IV				$\frac{RMSFE^{h=4}_{70-83}}{RMSFE^{h=4}_{84-04}}$
	h=1	h=2	h=4	h=8	h=1	h=2	h=4	h=8	
AR(AIC) RMSFE	1.72	1.75	1.89	2.38	0.78	0.68	0.62	0.73	
Relative MSFEs									
AR(AIC)	1.00	1.00	1.00	1.00	1.00	**1.00**	1.00	1.00	.33
AO	1.95	1.57	1.06	1.00	1.22	1.10	**0.89**	**0.84**	.30
PC-u	**0.85**	0.92	0.88	**0.61**	**0.95**	1.11	1.48	1.78	.42
PC-Δu	0.87	**0.87**	**0.86**	0.64	1.06	1.27	1.83	2.21	.48
PC-$u-\bar{u}^{1-Sided}$	0.88	0.99	0.98	0.87	1.06	1.29	1.84	2.39	.45
PC-Permits	0.93	1.02	0.98	0.78	1.08	1.23	1.31	1.52	.38

B. PCE Deflator (Core)

	1970:I – 1983:IV				1984:I – 2004:IV				$\frac{RMSFE^{h=4}_{70-83}}{RMSFE^{h=4}_{84-04}}$
	h=1	h=2	h=4	h=8	h=1	h=2	h=4	h=8	
AR(AIC) RMSFE	1.37	1.46	1.66	2.11	0.68	0.60	0.56	0.62	
Relative MSFEs									
AR(AIC)	1.00	1.00	1.00	1.00	**1.00**	**1.00**	1.00	1.00	.33
AO	2.55	1.90	1.14	0.99	1.59	1.11	**0.71**	**0.59**	.26
PC u	1.08	0.99	0.94	**0.60**	1.03	1.04	1.32	1.58	.40
PC Δu	**0.98**	**0.90**	**0.83**	0.65	1.06	1.14	1.45	1.95	.44
PC $u-\bar{u}^{1-Sided}$	1.07	0.97	1.16	1.00	1.07	1.16	1.51	2.05	.38
PC Permits	1.21	1.08	1.28	1.03	1.01	1.11	1.18	1.31	.32

C. PCE Deflator (All Items)

	1970:I – 1983:IV				1984:I – 2004:IV				$\frac{RMSFE^{h=4}_{70-83}}{RMSFE^{h=4}_{84-04}}$
	h=1	h=2	h=4	h=8	h=1	h=2	h=4	h=8	
AR(AIC) RMSFE	1.71	1.76	2.13	2.78	1.04	0.94	0.90	0.92	
Relative MSFEs									
AR(AIC)	1.00	1.00	1.00	1.00	1.00	**1.00**	1.00	1.00	.42
AO	2.43	1.98	1.13	0.97	1.57	1.18	**0.74**	**0.79**	.34
PC-u	0.78	0.92	0.99	0.69	**0.92**	1.04	1.26	1.77	.47
PC-Δu	**0.74**	**0.83**	**0.87**	**0.68**	0.95	1.09	1.28	1.95	.51
PC-$u-\bar{u}^{1-Sided}$	0.78	0.89	0.97	0.86	0.96	1.12	1.40	2.29	.50
PC-Permits	1.01	0.91	0.96	0.81	1.02	1.06	0.94	1.20	.42

Notes to Table 1: The first row of entries are root mean squared forecast errors (RMSFEs) of the AR(AIC) benchmark forecast. For the remaining rows, the first eight numerical columns report the MSFE of the forecasting model, relative to the AR(AIC) benchmark (hence AR(AIC) = 1.00). The final column reports the reduction in RMSFE from the 1970-1983 period to the 1984-2004 period for the row forecasting method, at the four-quarter horizon. Bold entries denote the lowest MSFE for that period/horizon. All forecasts are pseudo out-of-sample.

Table 2: Autocorrelations of $\Delta\pi_t$

	GDP		PCE (Core)		PCE (All)	
Lag	**1970:I – 1983:IV**	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**
1	-0.187 (0.102)	**-0.416** (0.109)	**-0.263** (0.102)	**-0.396** (0.109)	**-0.220** (0.102)	**-0.382** (0.109)
2	-0.148 (0.106)	-0.084 (0.127)	0.094 (0.109)	-0.143 (0.125)	-0.090 (0.107)	-0.174 (0.124)
3	-0.006 (0.108)	-0.117 (0.127)	-0.087 (0.110)	0.064 (0.127)	0.167 (0.108)	0.143 (0.127)
4	0.150 (0.108)	**0.395** (0.129)	0.022 (0.110)	0.047 (0.127)	0.048 (0.110)	-0.023 (0.129)
5	-0.048 (0.110)	-0.268 (0.142)	0.064 (0.110)	-0.082 (0.128)	-0.130 (0.111)	-0.115 (0.129)
6	-0.011 (0.110)	-0.020 (0.148)	-0.101 (0.111)	0.009 (0.128)	0.050 (0.112)	0.021 (0.130)
7	-0.062 (0.110)	-0.000 (0.148)	-0.026 (0.112)	0.165 (0.128)	0.024 (0.112)	0.242 (0.130)
8	0.001 (0.110)	0.304 (0.148)	-0.127 (0.112)	-0.158 (0.130)	-0.212 (0.112)	-0.131 (0.135)
$\hat{\sigma}_{\Delta\pi}$	1.30	0.91	1.10	0.73	1.32	1.15

Notes: Entries are autocorrelations for the indicated sample period, with standard errors in parentheses. Bold entries are significant at the 5% (two-sided) significance level. The row labeled $\hat{\sigma}_{\Delta\pi}$ reports the standard deviation of the quarterly change of inflation (at an annual rate).

Table 3.
A Measure of Inflation Persistence

	GDPD		PCE-core		PCE-all	
	1970:I – 1983:IV	1984:I – 2004:IV	1970:I – 1983:IV	1984:I – 2004:IV	1970:I – 1983:IV	1984:I – 2004:IV
Largest AR Root:: 90% conf. interval	0.884 - 1.030	0.852 - 1.032	0.879 - 1.029	0.913 - 1.040	0.903 - 1.032	0.834 - 1.029

Notes: The 90% confidence interval for the largest autoregressive root of inflation was computed using Stock's (1991) method of inverting the augmented Dickey-Fuller (ADF) test statistic. The ADF statistic was computed with AIC lag length choice and a constant (no time trend).

Table 4
IMA(1,1) Models of $\Delta\pi_t$ and their Unobserved Component Representations

	GDPD		PCE-core		PCE-all	
	1970:I − 1983:IV	1984:I − 2004:IV	1970:I − 1983:IV	1984:I − 2004:IV	1970:I − 1983:IV	1984:I − 2004:IV
IMA parameters: $\Delta\pi_t = (1 - \theta B)a_t$						
θ	0.275 (.085)	0.656 (.088)	0.252 (.063)	0.677 (.094)	0.249 (.094)	0.688 (.088)
σ_a	1.261 (.070)	0.753 (.070)	1.053 (.058)	0.604 (.055)	1.273 (.080)	0.966 (.073)
UC parameters						
σ_ε	0.914 (.118)	0.259 (.072)	0.787 (.079)	0.195 (.063)	0.957 (.121)	0.301 (.093)
σ_η	0.662 (.110)	0.610 (.068)	0.529 (.074)	0.497 (.051)	0.635 (.080)	0.801 (.072)
p-values of Wald Tests of IMA(1,1) specification against higher-order alternatives						
IMA(1,1) vs. ARIMA(1,1,1)	0.42	0.98	0.32	0.33	0.98	0.91
IMA(1,1) vs. IMA(1,4)	0.40	0.13	0.66	0.73	.004	0.46
ARIMA(1,0,1) parameters: $(1 - \phi B)\pi_t = (1 - \theta B)a_t$						
ϕ	0.987 (.018)	0.989 (.011)	0.990 (.017)	0.992 (.008)	0.986 (.019)	0.992 (.013)
θ	−0.261 (.102)	−0.673 (.084)	−0.243 (.105)	−0.679 (.083)	−0.240 (.104)	−0.687 (.083)
Tests for parameter stability in the UC model						
t-statistic for $\sigma_{\varepsilon,70\text{-}83} = \sigma_{\varepsilon,84\text{-}04}$ (*p*-value)	–	–4.75 (<.001)	–	–5.89 (<.001)	–	–4.31 (<.001)
t-statistic for $\sigma_{\eta,70\text{-}83} = \sigma_{\eta,84\text{-}04}$ (*p*-value)	–	–0.41 (.684)	–	–0.35 (.727)	–	1.08 (.278)
QLR (*p*-value)	–	31.99 (<.01)	–	24.84 (<.01)	–	17.69 (<.01)
Variance decomposition of four-quarter inflation forecasts: UC model						
4-quarter MSE	1.99	0.35	1.44	0.21	2.12	0.53
MSE due to:						
filtering error	0.32	0.13	0.21	0.08	0.30	0.20
trend shocks	1.57	0.13	1.16	0.07	1.72	0.17
transitory shocks	0.11	0.09	0.07	0.06	0.10	0.16

Notes to table 4: The first block reports estimated parameters of the IMA(1,1) model, the second block reports the corresponding parameters of the unobserved components model, the third block reports tests of the IMA(1,1) specification against ARIMA models with more parameters, the fourth block reports estimates of ARIMA(1,0,1) models that do not impose a unit root in inflation; in all these blocks, standard errors are in parentheses. The fifth block reports tests for parameter stability, first one parameter at a time in the UC model with an imposed break in 1984 (shown as *t*-statistics), then the Quandt (maximal) likelihood ratio (QLR) statistic (df = 2) over all break dates in the inner 70% of the sample (the QLR statistic). The *p*-values in parentheses in the final block take the 1984 break date as exogenous for the first two rows, but the QLR critical values allow for an endogenous break (Andrews (1993)). The final block reports a decomposition of the total four-quarter ahead forecast error variance (first row), based on the UC model, into the three components of filtering (signal extraction) error, future permanent disturbances, and future transitory disturbances.

Table 5. Pseudo Out-of-Sample Forecasting Performance of Additional Univariate Models: MSFEs, Relative to AR(AIC)

(a) GDP Deflator

	1970:I – 1983:IV				1984:I – 2004:IV			
Model	**h=1**	**h=2**	**h=4**	**h=8**	**h=1**	**h=2**	**h=4**	**h=8**
Recursive forecasts								
AR(AIC)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AO	1.95	1.57	1.06	1.00	1.22	1.10	**0.89**	**0.84**
MA(1)	**0.82**	0.83	0.87	0.89	1.01	1.03	0.98	0.89
AR(4)	0.95	1.08	1.05	0.93	0.93	0.96	0.99	0.94
Rolling forecasts								
AR(AIC)	0.97	1.05	0.99	0.83	0.95	1.08	1.17	1.18
AR(4)	0.98	1.15	1.06	0.94	**0.92**	0.95	1.06	1.04
MA(1)	**0.82**	**0.82**	**0.86**	**0.88**	0.99	0.98	0.93	0.87
Nelson-Schwert								
NS77 MA(2)	0.88	0.91	0.95	0.89	0.93	**0.91**	0.92	0.87
Fixed-parameter comparisons								
UC-SV, $\gamma = 0.2$	0.77	0.79	0.82	0.88	0.96	0.94	0.90	0.83
MA(1) $\theta = 0.25$	0.79	0.80	0.82	0.87	1.05	1.11	1.05	0.93
MA(1) $\theta = 0.65$	0.97	0.94	0.96	0.90	0.90	0.87	0.89	0.82

(b) PCE-Core

	1970:I – 1983:IV				1984:I – 2004:IV			
Model	**h=1**	**h=2**	**h=4**	**h=8**	**h=1**	**h=2**	**h=4**	**h=8**
Recursive forecasts								
AR(AIC)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AO	2.55	1.90	1.14	0.99	1.59	1.11	**0.71**	**0.59**
MA(1)	0.97	0.85	0.89	0.90	0.94	0.92	0.89	0.70
AR(4)	1.11	1.18	1.13	0.98	1.01	1.02	0.97	0.87
Rolling forecasts								
AR(AIC)	1.10	1.12	0.99	1.12	1.21	1.08	1.05	0.76
AR(4)	1.16	1.24	1.17	1.03	1.22	1.16	1.05	0.84
MA(1)	**0.96**	**0.83**	**0.87**	0.89	0.89	0.84	0.81	0.68
Nelson-Schwert								
NS77 MA(2)	1.08	1.01	0.99	**0.86**	**0.83**	**0.78**	0.75	0.65
Fixed-parameter comparisons								
UC-SV, $\gamma = 0.2$	0.93	0.84	0.88	0.88	0.91	0.84	0.78	0.67
MA(1) $\theta = 0.25$	0.92	0.82	0.85	0.86	0.96	0.95	0.92	0.73
MA(1) $\theta = 0.65$	1.16	1.05	1.01	0.88	0.81	0.75	0.71	0.61

Table 5, continued

(c) PCE-All

	1970:I − 1983:IV				1984:I − 2004:IV			
Model	**h=1**	**h=2**	**h=4**	**h=8**	**h=1**	**h=2**	**h=4**	**h=8**
Recursive forecasts								
AR(AIC)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AO	2.43	1.98	1.13	0.97	1.57	1.18	0.74	0.79
MA(1)	0.87	0.89	0.84	**0.83**	0.98	0.94	0.90	0.87
AR(4)	0.99	1.09	1.05	1.00	0.99	0.99	1.00	0.94
Rolling forecasts								
AR(AIC)	1.03	0.96	0.96	0.97	1.04	1.00	0.94	0.96
AR(4)	1.01	1.13	1.08	1.08	1.04	1.01	0.98	0.97
MA(1)	**0.86**	**0.88**	**0.82**	**0.83**	0.93	0.86	0.80	0.83
Nelson-Schwert								
NS77 MA(2)	1.03	1.10	1.01	0.85	**0.87**	**0.79**	**0.73**	**0.73**
Fixed-parameter comparisons								
UC-SV, $\gamma = 0.2$	0.85	0.87	0.80	0.81	0.97	0.92	0.87	0.84
MA(1) $\theta = 0.25$	0.83	0.83	0.79	0.80	1.01	0.99	0.94	0.90
MA(1) $\theta = 0.65$	1.13	1.13	1.03	0.87	0.86	0.77	0.72	0.73

Notes: Entries are MSFEs, relative to the recursively estimated AR(AIC). Bold entries are the smallest relative MSFE for the indicated series/period/horizon, among the out-of-sample (NS77) and pseudo out-of-sample forecasts. The fixed-parameter models do not generate pseudo out-of-sample forecasts because their parameters are not estimated using recursive or rolling samples.

Table 6.
In-Sample Summary Statistics for Bivariate ($\Delta\pi_t$, x_t) Processes for Various Activity Measures *x*

Activity variable	1960:I-1983:IV		1984:I-2004:IV	
	Relative MSE, VAR v. univariate, $h = 4$	Coherence($\Delta\pi_t$, x_t) at bus. cycle frequencies	Relative MSE, VAR v. univariate, $h = 4$	Coherence($\Delta\pi_t$, x_t) at bus. cycle frequencies
(a) GDP deflator				
Δu	0.72	0.50	0.92	0.52
$\Delta lnGDP$	0.82	0.43	0.87	0.41
ln(Permits)	0.73	0.54	0.82	0.53
capacity util.	0.66	0.54	0.90	0.47
ugap(2s)	0.76	0.48	0.88	0.58
ygap(2s)	0.77	0.34	0.83	0.41
(b) PCE-core				
Δu	0.80	0.53	0.97	0.34
$\Delta lnGDP$	0.90	0.39	0.98	0.37
ln(Permits)	0.87	0.51	0.81	0.50
capacity util.	0.68	0.61	0.96	0.39
ugap(2s)	0.75	0.55	0.87	0.40
ygap(2s)	0.72	0.34	0.90	0.37
(c) PCE-all				
Δu	0.68	0.60	0.94	0.46
$\Delta lnGDP$	0.69	0.49	0.92	0.43
ln(Permits)	0.73	0.65	0.78	0.67
capacity util.	0.78	0.59	0.93	0.48
ugap(2s)	0.86	0.59	0.91	0.58
ygap(2s)	0.86	0.42	0.91	0.40

Notes: The relative MSE row reports the implied mean squared error for 4-quarter ahead predictions from the iterated VAR(4), relative to the mean squared error of the iterated 4-quarter ahead predictions from the iterated univariate ARIMA(8,1,4) model for inflation implied by the VAR(4). The coherence is the average coherence over the business cycle frequencies corresponding to periods between 6 and 24 quarters. All entries are computed from a bivariate VAR(4) estimated for the ($\Delta\pi_t$, x_t) pair over the indicated subsample, where the activity variable x_t is listed in the first column.

Table 7. In-sample Four-quarter Ahead ADL(4,4) Models of Inflation Using Two-sided Gaps (*xgap*):

$$\pi^4_{t+4} - \pi_t = \mu + \beta xgap_{t-1} + \delta(\mathrm{B})\Delta xgap_{t-1} + \alpha(\mathrm{B})\Delta\pi_{-1} + u^4_t$$

(a) *x* = unemployment gap (2-sided)

	GDPD		PCE-Core		PCE-All	
	1970:I – 1983:IV	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**
relative MSE, ADL (4,4) v. AR(4)	0.641	0.817	0.686	0.771	0.682	0.882
gap coefficient β (SE)	-0.194 (0.150)	-0.225 (0.074)	-0.366 (0.128)	-0.249 (0.069)	0.100 (0.181)	-0.168 (0.090)
Granger causality F-test: (*p*-value)	0.000	0.004	0.000	0.001	0.000	0.084
Chow test for a break in coefficients on **(*p*-value):**						
xgap		0.856		0.420		0.186
xgap and Δx_t lags		0.002		0.011		0.000
$\Delta\pi$ lags		0.563		0.065		0.112
all coefficients		0.000		0.000		0.000

(b) *x* = GDP gap (2-sided)

	GDPD		PCE-Core		PCE-All	
	1970:I – 1983:IV	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**
relative MSE, ADL (4,4) v. AR(4)	0.689	0.778	0.670	0.798	0.652	0.852
gap coefficient (SE)	0.159 (0.056)	0.118 (0.035)	0.209 (0.047)	0.130 (0.035)	0.054 (0.070)	0.104 (0.045)
Granger causality F-test: (*p*-value)	0.043	0.000	0.028	0.002	0.000	0.004
Chow test for a break in coefficients on **(*p*-value):**						
x-gap		0.539		0.172		0.541
x-gap and Δx_t lags		0.098		0.004		0.001
$\Delta\pi$ lags		0.647		0.017		0.058
all coefficients		0.029		0.000		0.000

(c) x = Δu (gap term excluded)

	GDPD		PCE-Core		PCE-All	
	1970:I – 1983:IV	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**
relative MSE, ADL (4,4) v. AR(4)	0.652	0.896	0.736	0.923	0.684	0.907
gap coefficient (SE)	–	–	–	–	–	–
Granger causality F-test: (*p*-value)	0.000	0.019	0.001	0.119	0.000	0.162
Chow test for a break in coefficients on **(*p*-value):.**						
Δx_t lags		0.000		0.003		0.000
$\Delta\pi$ lags		0.653		0.068		0.183
all coefficients		0.000		0.000		0.000

Notes to table 7: Gap variables are created as deviations from a two-sided centered MA(80) low pass filter, with pass band of periodicities greater than 60 quarters. All entries computed from the coefficients of the four-quarter direct regression given in the table heading, estimated using the sample period in the column heading, with four lags of $\Delta\pi$ and 3 lags of Δx (which corresponds to four lags of x). The relative MSE in the first row is the implied in-sample MSE of the ADL forecast, relative to the AR(4) forecast.

Table 8
Four-quarter Ahead Forecast Error Variance Decomposition from Projections of Univariate Forecast Errors onto Various Activity Measures *x*

	GDP deflator			PCE–Core			PCE-All		
	1970:I – 1983:IV	1984:I – 2004:IV	diff	1970:I – 1983:IV	1984:I – 2004:IV	diff	1970:I – 1983:IV	1984:I – 2004:IV	diff
4-qtr forecast error mean SS (MSS)	1.86	0.33	1.53	1.45	0.20	1.26	2.13	0.55	1.57
x = unemployment gap (2-sided gap)									
error MSS	1.27	0.26	1.01	1.06	0.15	0.90	1.44	0.48	0.96
explained MSS	0.59	0.06	0.53	0.40	0.04	0.35	0.68	0.07	0.61
Δ(X′X)			0.21			0.14			0.22
Δ(coefficients)			0.31			0.22			0.39
x = GDP gap (2-sided gap)									
error MSS	1.36	0.25	1.11	1.02	0.16	0.85	1.37	0.46	0.90
explained MSS	0.50	0.08	0.42	0.44	0.03	0.40	0.76	0.09	0.67
Δ(X′X)			0.28			0.18			0.42
Δ(coefficients)			0.14			0.22			0.26
x = Δu (gap term excluded)									
error MSS	1.27	0.30	0.97	1.09	0.18	0.91	1.46	0.50	0.96
explained MSS	0.59	0.03	0.56	0.36	0.02	0.35	0.67	0.05	0.61
Δ(X′X)			0.18			0.10			0.23
Δ(coefficients)			0.38			0.25			0.38

Notes: Entries are the mean sum of squared 4-quarter ahead forecast errors constructed from regressions of period-specific univariate IMA(1,1) forecast errors onto 4 lags of the activity variable and a constant. Entries are the in-sample mean of error sum of squares, the explained (regression) sum of squares, and the decomposition of the change in the explained sum of squares from the first period to the second into changes in the variance of x (denoted $\Delta(X'X)$) and the change in the regression coefficients (denoted Δ(coefficients)). Gaps are two-sided, computed as discussed in the notes to Table 7.

Table 9
Four-quarter Ahead Regressions of UC-SV Forecast Error on Activity Gaps

	GDP deflator		**PCE-core**		**PCE-all**	
	1970:I – 1983:IV	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**	**1970:I – 1983:IV**	**1984:I – 2004:IV**
(a) unemployment gap (2-sided)						
β	-0.40 (.14)	-0.16 (.09)	-0.51 (.14)	-0.21 (.09)	-0.32 (.17)	-0.03 (.12)
SER	1.62	0.57	1.38	0.45	1.80	0.85
$\bar{R}^2$	0.08	0.04	0.17	.11	0.03	-0.01
(b) GDP gap (2-sided)						
β	0.24 (.08)	0.12 (.04)	0.27 (.07)	0.10 (.04)	0.24 (.09)	0.04 (.05)
SER	1.55	0.55	1.32	0.46	1.70	0.84
$\bar{R}^2$	0.15	0.10	0.25	0.10	0.13	-0.01

Table 10
Relative MSFEs of Rolling Multivariate Four-quarter Ahead Pseudo Out-of-Sample Forecasts, 1984 – 2004

Dependent variable: four-quarter ahead forecast error from UC-SV ($\gamma = .2$) model
Regressors: constant, indicated number of lags of the activity variable x_t, and (as indicated) $\Delta\pi_t$

Activity variable	x lag length	GDPD		PCE-core		PCE-all	
		const	const, $\Delta\pi_t$	const	const, $\Delta\pi_t$	const	const, $\Delta\pi_t$
Δu	1	1.19	1.14	1.09	1.07	1.06	1.02
	BIC	2.04	2.11	1.13	1.07	1.36	1.34
unemployment gap (1-sided)	1	1.19	1.19	1.08	1.07	1.05	1.01
	BIC	1.19	1.18	1.08	1.07	1.31	1.29
ΔlnGDP	1	1.03	1.02	1.04	1.04	1.02	0.97
	BIC	1.67	1.73	1.50	1.51	1.31	1.29
GDP gap (1-sided)	1	1.24	1.23	1.11	1.11	1.12	1.08
	BIC	1.29	1.28	1.11	1.11	1.27	1.23
ln(Permits)	1	1.45	1.47	1.16	1.17	1.11	1.09
	BIC	1.45	1.47	1.21	1.15	1.12	1.12
Δln(Permits)	1	1.05	1.04	1.04	1.02	1.01	0.96
	BIC	1.06	1.05	1.04	1.02	1.02	0.97
capacity utilization	1	1.00	0.97	1.24	1.26	1.03	0.98
	BIC	1.02	1.04	1.33	1.32	1.04	1.02
Δcapacity utilization	1	1.11	1.07	1.06	1.05	1.05	1.01
	BIC	1.64	1.66	1.07	1.06	1.29	1.28

Notes: Relative MSFEs are relative to the univariate UC-SV forecast (which corresponds to a forecast of zero in the four-quarter ahead ADL model). Rolling estimates are computed using a 40-quarter window that ends at the forecast date. BIC lag lengths are selected with lags $1 \le p \le 4$. The column heading “const” and “const, $\Delta\pi_t$” indicate the regressors included in the direct ADL regression, in addition to the indicated number of lags of the x variable.

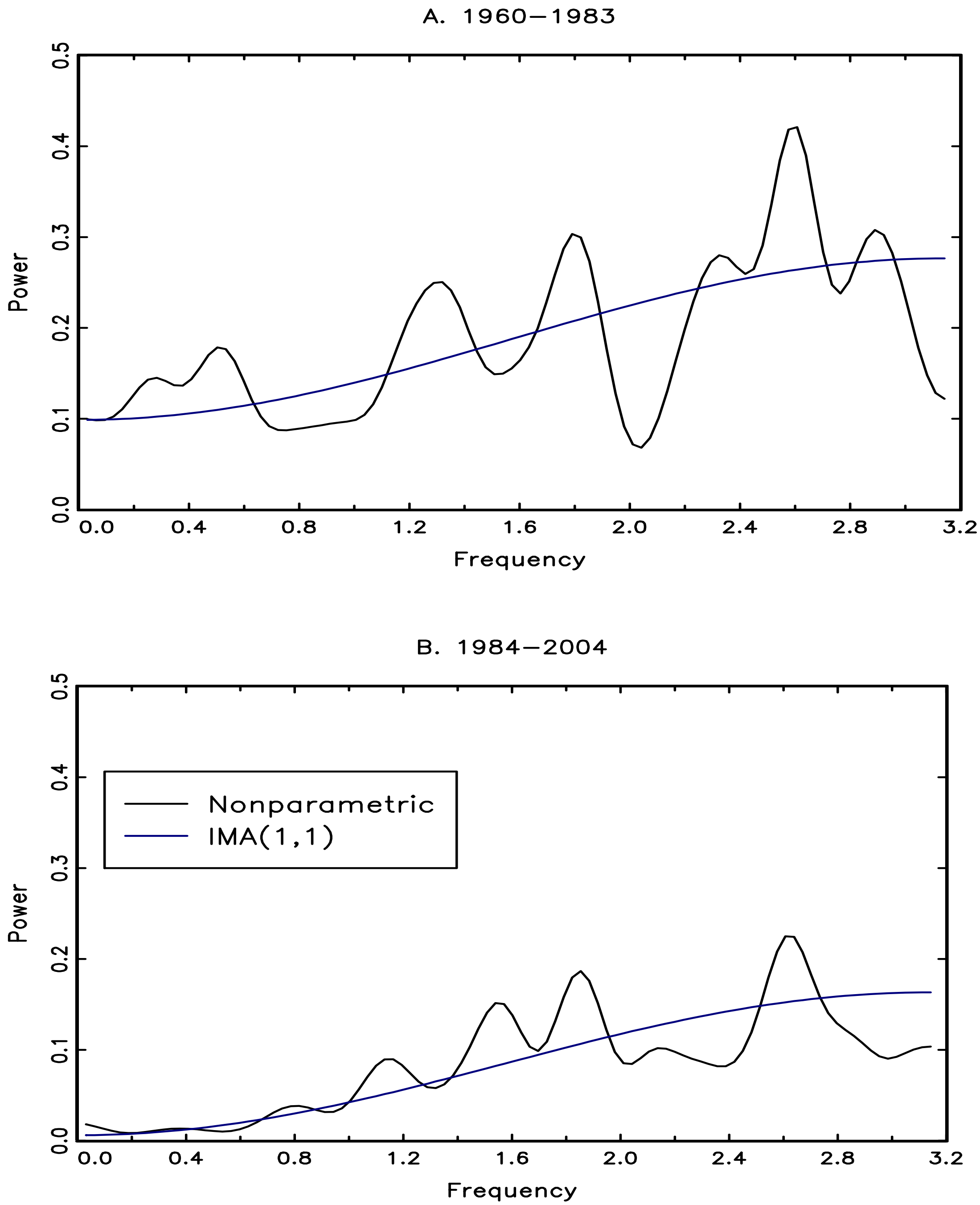


Figure 1. Parametric and Nonparametric Estimates of the Spectrum of the first difference of PCE(core) inflation, 1970-1983 and 1984-2004.

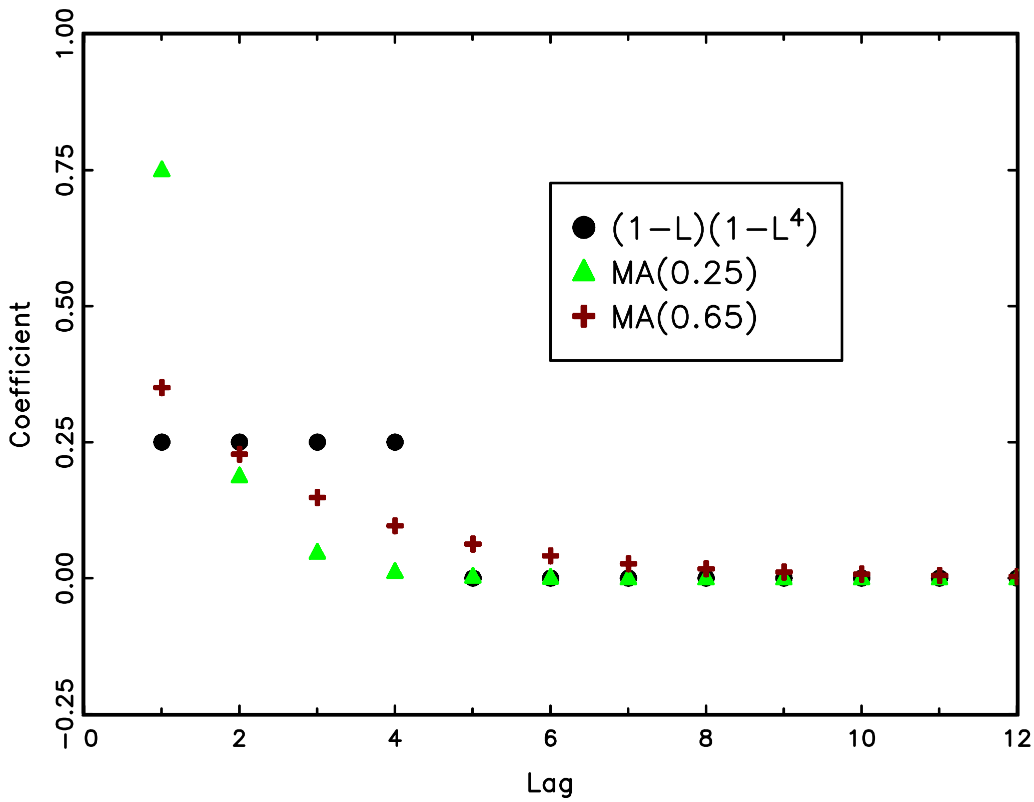
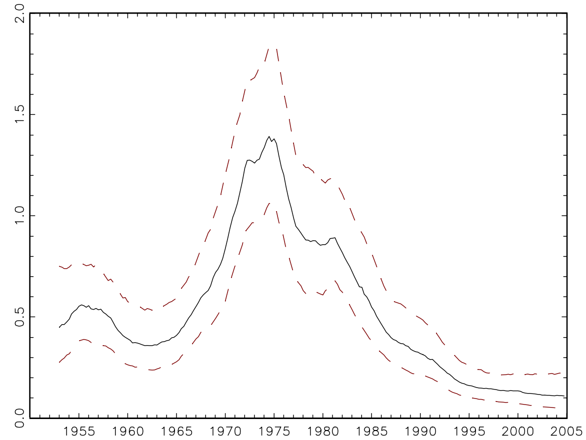
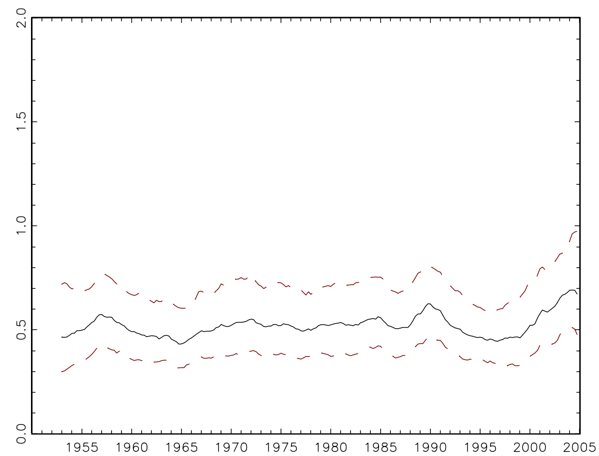


Figure 2. Implied forecast weights on lagged quarterly inflation for forecasts of four-quarter inflation computed using the Atkeson-Ohanian model $(1 - L)(1 - L^4)$ and using an IMA(1,1) model with $\theta = .25$ and $\theta = .65$.

(a) Standard deviation of permanent innovation, $\sigma_{\varepsilon,t}$



(b) Standard deviation of transitory innovation, $\sigma_{\eta,t}$



(c) Implied IMA(1,1) coefficient θ

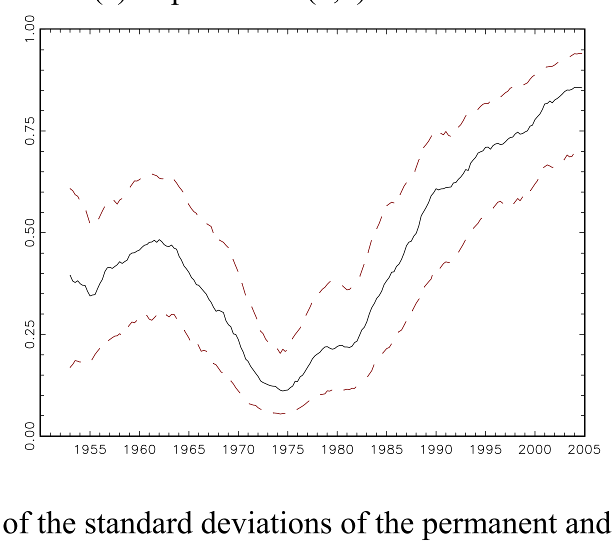
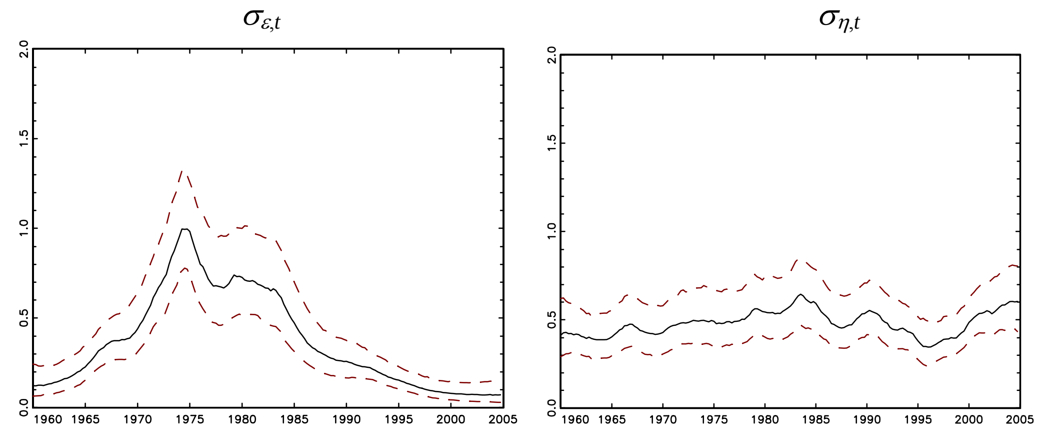


Figure 3. Estimates of the standard deviations of the permanent and transitory innovations, and of the implied IMA(1,1) coefficient, using the TC-SV(.2) model: 16.5%, 50%, and 83.5% quantiles of the posterior distributions, GDP deflator, 1953-2004

(a) PCE (Core)



(b) PCE (All Items)

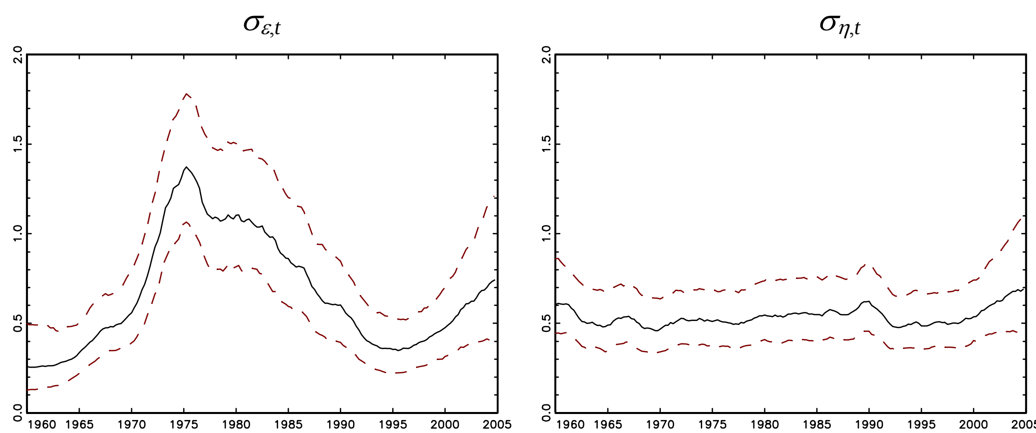


Figure 4. Estimates of the standard deviations of the permanent and transitory innovations, and of the implied IMA(1,1) coefficient, using the TC-SV(.2) model: 16.5%, 50%, and 83.5% quantiles of the posterior distributions, 1959-2004.

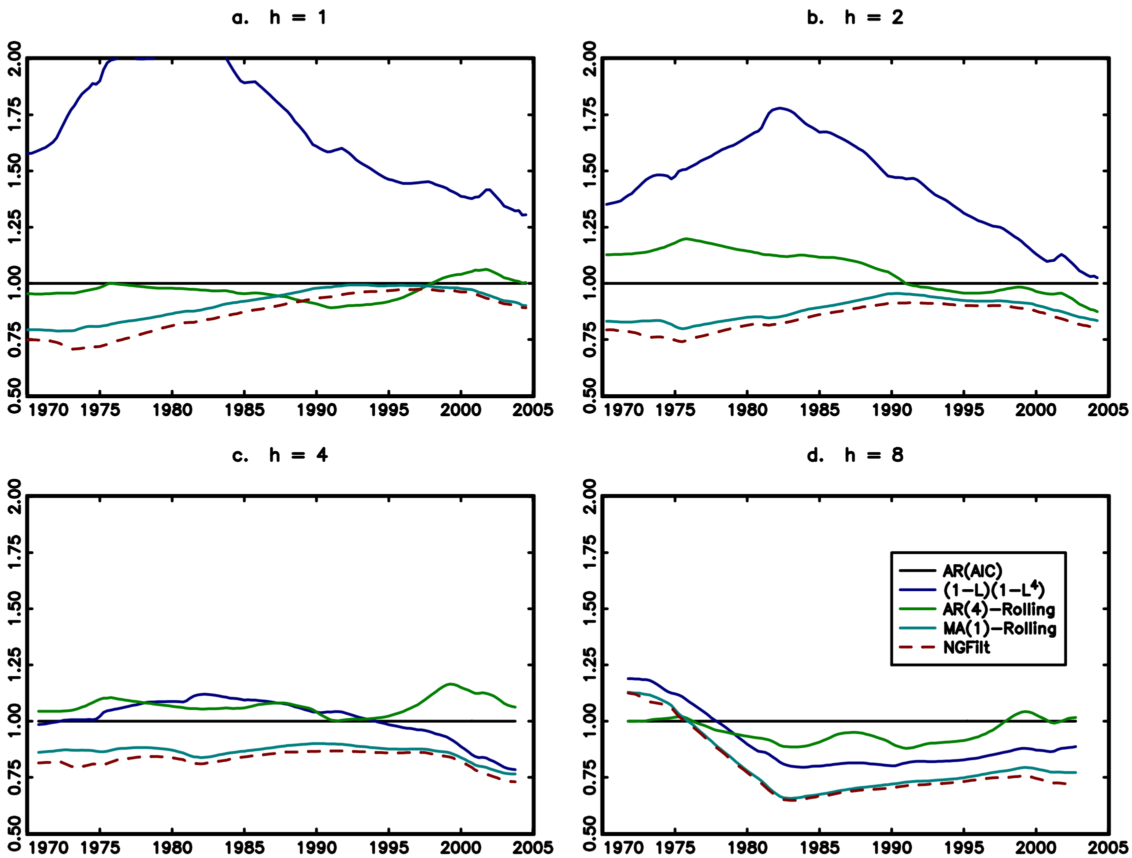


Figure 5. Smoothed relative mean squared forecast errors of various forecasts, relative to the recursive AR(AIC) benchmark: GDP deflator.

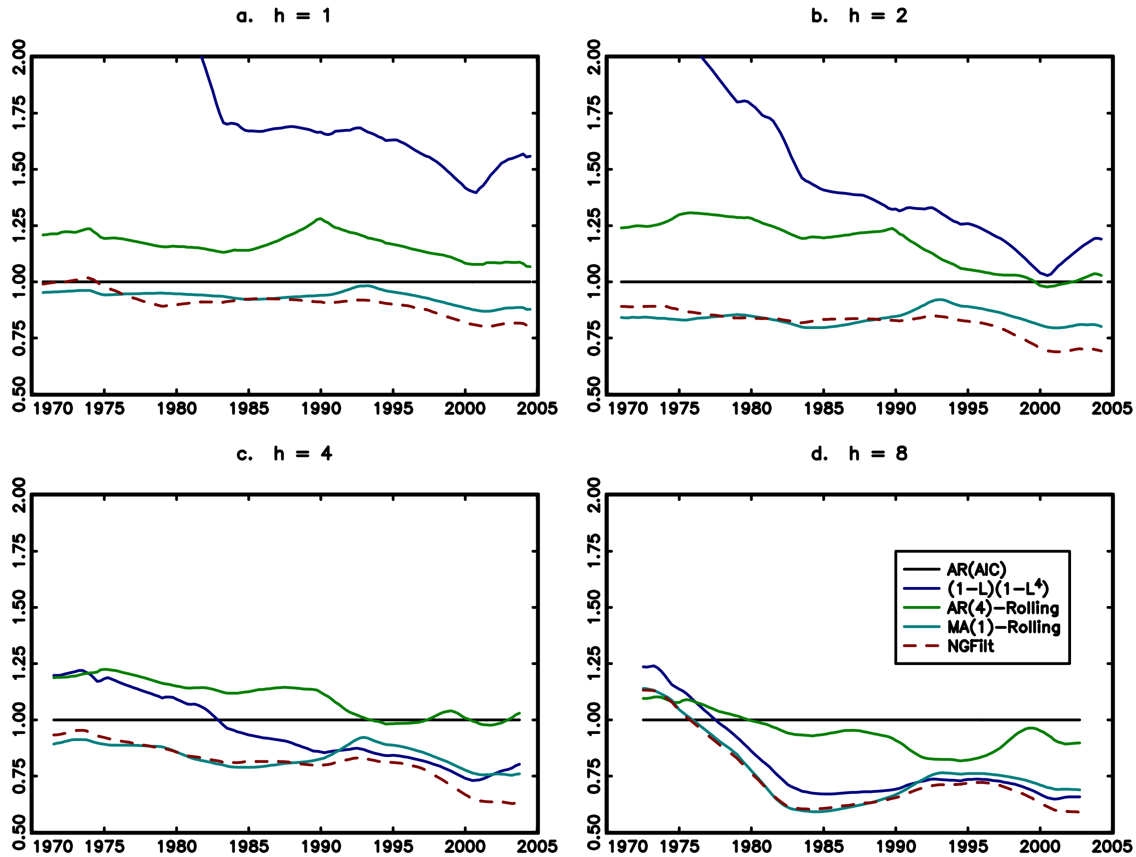


Figure 6. Smoothed relative mean squared forecast errors of various forecasts, relative to the recursive AR(AIC) benchmark: Core PCE

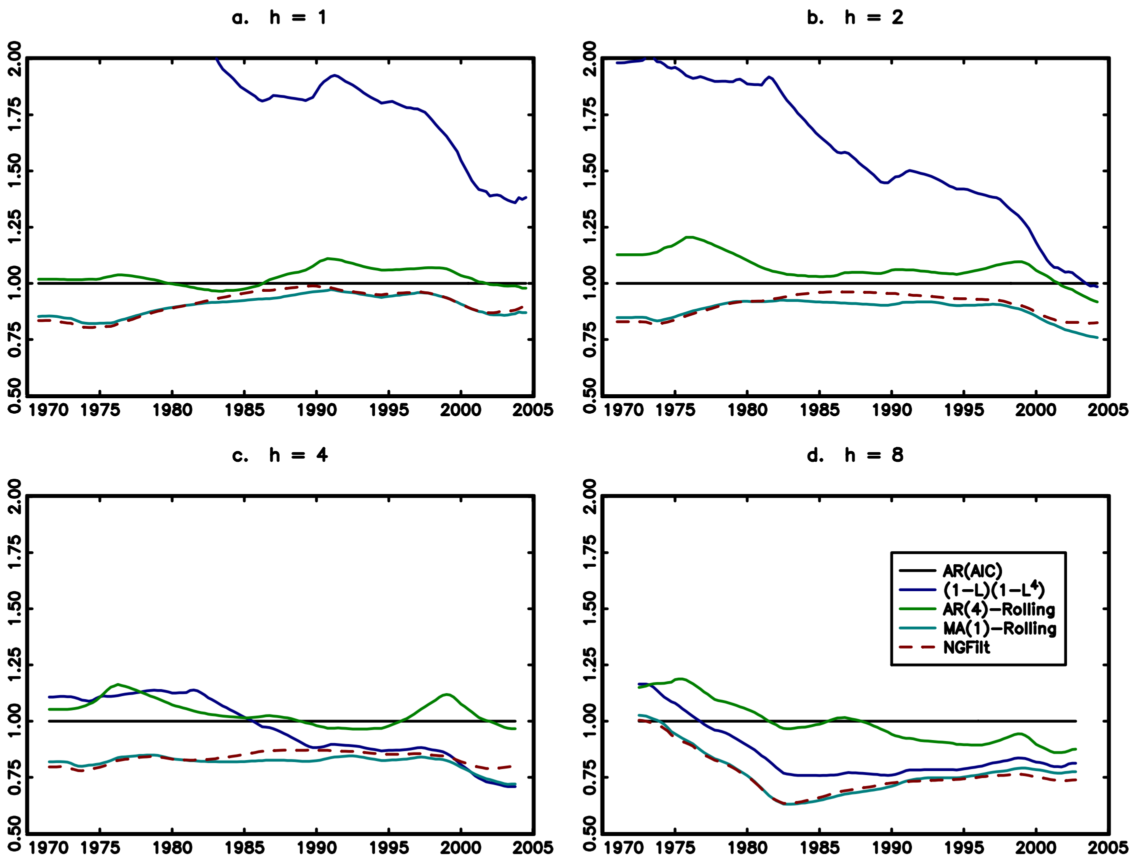


Figure 7. Smoothed relative mean squared forecast errors of various forecasts, relative to the recursive AR(AIC) benchmark: PCE-all

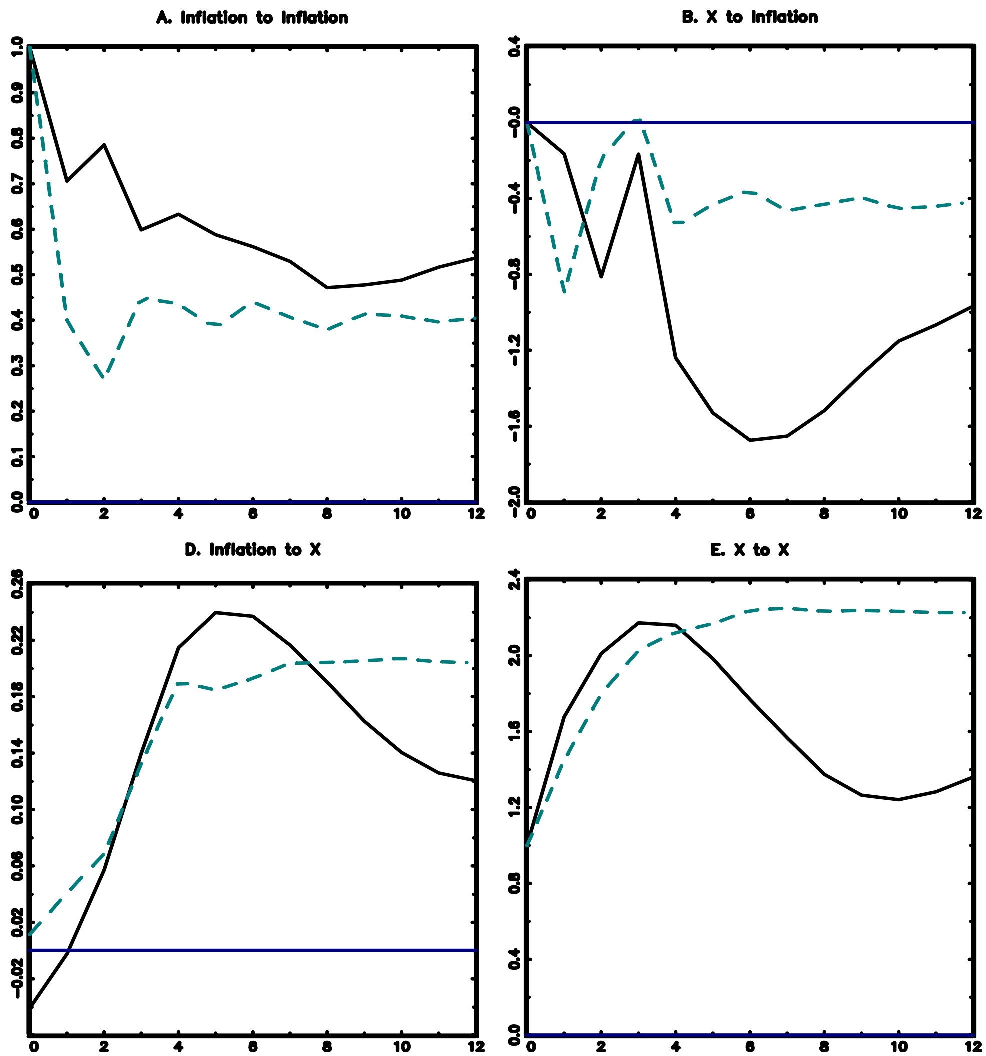


Figure 8. Impulse response functions from a VAR(4) for ($\Delta\pi_t$, Δu_t) (Cholesky decomposition, $\Delta\pi_t$ ordered first): 1960-1983 (solid line) and 1984 – 2004 (dashed line)

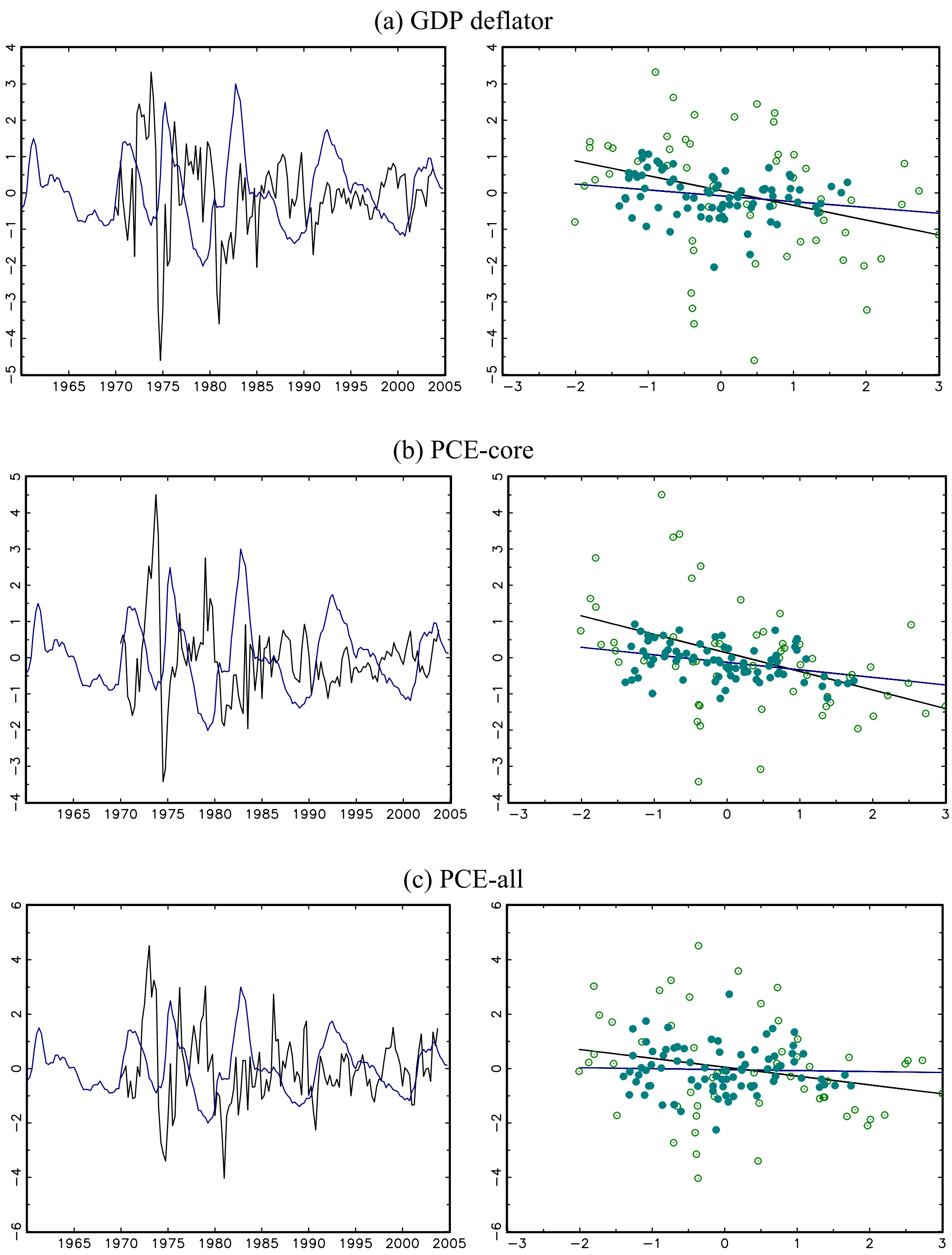


Figure 9. Time series plot (left) and scatterplot of residual (right) from UC-SV model vs. the 2-sided unemployment gap. Scatterplot: light green, 1970–83; dark green, 1984–04.