# (Un)Predictability and Macroeconomic Stability*

Antonello D'Agostino
Central Bank and Financial Services Authority of Ireland
Domenico Giannone
European Central Bank, ECARES and CEPR
Paolo Surico
Bank of England

January 2009

**Abstract**

The ability of popular statistical methods, the Federal Reserve Greenbook and the Survey of Professional Forecasters to improve upon naive forecasts of inflation and real activity has declined significantly in U.S. data moving from the pre- to the post-1985 sample. The decline is larger for institutional forecasters and models based on large information sets. In the most recent period, there is evidence of predictability for inflation only one month ahead, and for unemployment rate and nonfarm payrolls at most horizons. Counterfactual analyses suggest that a change in the estimated coefficients has been relatively more important than a change in the estimated error variances to explain these findings.

JEL Classification: E37, E47, C22, C53.

Keywords: predictability, forecasting models, Fed Greenbook, Survey of Professional Forecasts.

*We are grateful to Raffaella Giacomini, Jan Groen, Christoph Schleicher, Mark Watson and Karl Whelan for helpful discussions, and seminar participants at the North American meeting of the Econometric Society 2007, the meeting of the European Economic Association 2007, Trinity College, and Forecasting in Rio 2008 for comments. A previus version of this paper has circulated as ECB working paper No. 605. The views expressed in this paper are those of the authors, and do not necessarily reflect those of the Bank of England, the Central Bank and Financial Services Authority of Ireland or the European Central Bank. Address for correspondence - Antonello D'Agostino: Central Bank and Financial Services Authority of Ireland - Economic Analysis and Research Department, PO Box 559 - Dame Street, Dublin 2, Ireland. E-mail: antonello.dagostino@centralbank.ie; Domenico Giannone: European Central Bank, Directorate General Research, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany. E-mail: domenico.giannone@ecb.int; Paolo Surico: External Monetary Policy Committee Unit, Bank of England, Threadneedle street - EC2R 8AH - London, United Kingdom. E-mail: paolo.surico@bankofengland.co.uk

# 1 Introduction

The behavior of inflation and output in the United States has been characterized by two major episodes over the postwar history. The first episode was a period of large volatility that extended from the early 1970s to the mid-1980s. The second episode, from the second half of the 1980s to the present, is associated with far more stable inflation and output. The historical decline in volatility, documented first by Kim and Nelson (1999), McConnell and Perez-Quiros (2000), Blanchard and Simon (2001) and Cogley and Sargent (2002), is often referred to as the 'Great Moderation' and appears to hold across a wide number of sectors and countries (see Stock and Watson, 2003a).

Earlier contributions have noticed that the recent macroeconomic stability has implied a significant reduction in the uncertainty of the forecasts based on naive models. In this paper, we want to investigate whether the changes in macroeconomic volatility has been associated with changes in the relative predictability between sophisticated and naive models.[1] A large empirical literature has identified the mid-80s as the beginning of the Great Moderation (see for instance Kim and Nelson, 1999, and McConnell and Perez-Quiros, 2000). As we are interested in studying the connection between changes in volatility and changes in predictability, it makes most sense then to split the sample in the mid-80s.

Our main results can be summarized as follows. First, before 1985, using large information sets was very helpful to predict inflation and real activity. Today, in contrast, sophisticated models have no significant advantage upon naive models. Second, the fall in predictability is a common feature of many forecasting models including those used by public and private institutions. In particular, the forecasts for output and inflation of the Federal Reserve's Greenbook and the Survey of Professional Forecasters (SPF) are significantly more accurate than a random walk *only* before 1985. After 1985, in contrast, the hypothesis of equal predictive accuracy between naive random walk forecasts and the predictions of those institutions is not rejected for all horizons but the current quarter. Third, the decline in predictability is far more pronounced for institutional forecasters and methods based on large information sets than for univariate specifications. The fact that larger models are associated with larger historical changes suggests that the main source of change in predictability are the dynamic correlations between variables rather than the autocorrelations of output and inflation.

In Figure 1, the basic pattern that we uncover in the data is illustrated. In particular, we plot the relative predictability between 132 models and a naive forecast for the pre-1985 period against the relative predictability for the post-1985 period at one year horizon. Points below the 45 degree line indicates a decline in predictability. The figure suggests that after 1985, most forecast models have lost their relative advantage to predict consumer price index and industrial production. As for the 3-months treasury bills, there is some evidence of an improvement in predictability. We will explore these issues more formally in what follows.

The results of this paper may also be of interest for the empirical literature on asymmetric information. Romer and Romer (2000), for instance, consider a sample ending in the early 1990s and find that the Fed produced more accurate forecasts over inflation and output relative to several commercial providers. Our results imply that the informational advantage of the Fed and professional forecasters is, in fact, limited to the 1970s and the beginning of the 1980s. During the last two decades, in contrast, no forecast model has been better than *tossing a coin* beyond the first quarter horizon. This implies that, *on average*, uninformed or naive economic agents can effectively anticipate future macroeconomics developments. Econometric models and economists' judgement, however, are still helpful to forecast at the very short horizon.

The literature on forecasting methods, surveyed by Stock and Watson (2005), has devoted a great deal of attention towards identifying the best model for predicting inflation and output.

[1] Predictability is defined as the *ratio* between the Mean Squared Forecast Errors (MSFE) of the model of interest and the MSFE of a naive model (see Granger and Newbold, 1986, and Diebold and Kilian, 2001).

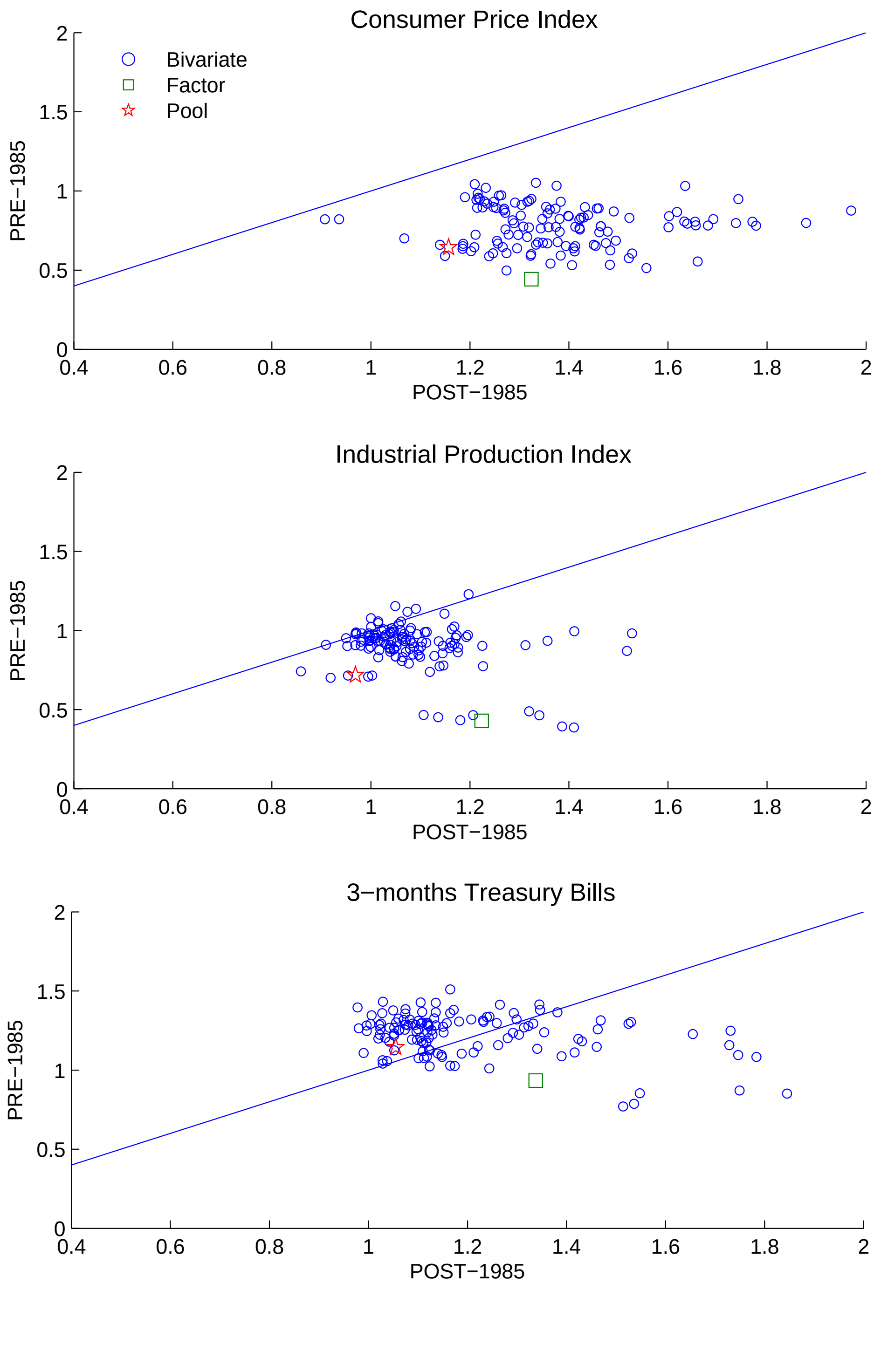


Figure 1: Relative predictability between 132 models and a naive forecast - pre-1985 period vs. post-1985 period - one year horizon.

The majority of studies, however, are based on full-sample periods. Our findings reveal that most of the full sample predictability of U.S. macroeconomic series comes indeed from the years before 1985. Long time series appear to assign a far larger weight to the earlier sub-sample, which is characterized by a larger volatility of inflation and output. The results presented here suggest that some caution should be used in evaluating the performance of alternative forecasting models on the basis of a pool of different sub-periods: parameter instability may affect full sample analyses.

The paper is organized as follows. The forecasting models are presented in Section 2. In Section 3 (4), we report the full sample (sub-sample) results. Section 5 shows that the models of the Fed and some commercial organizations are also associated with a remarkable fall in forecasting accuracy. In Section 6, we perform counterfactual analyses to shed some lights on the sources of the decline in macroeconomic predictability. The Appendix reports the definitions of the variables and the transformations applied.

# 2 The forecasting models

This section defines the concept of predictability and describes the data set. Our goal is to explore the nexus between the greater macroeconomic stability of the last two decades and the ability of several models to forecast inflation, real activity and interest rates. We construct forecasts for nine monthly key macroeconomic series: three price indices, four measures of real activity and two interest rates. The data set consists of monthly observations from 1959:1 through 2003:12 on 131 U.S. macroeconomic time series including also the nine variables of interest.

Forecasts are based on traditional univariate time series models as well as on models exploiting larger information. Using all variables as predictors poses, in fact, a serious curse of dimensionality problem for traditional models. Large cross-section forecasting methods, in contrast, can easily accommodate a large set of predictors. Among the latter, we consider two methods: factor model forecasts (employed by Stock and Watson, 2003b, and Giannone, Reichlin and Small, 2007); and pooling of forecasts (introduced by Bates and Granger, 1969). The first method is based on the notion that a few common factors can capture and describe most information in the data. The second method combines forecasts from small scale traditional time series models.

The three nominal variables are Producer Price Index ($PPI$), Consumer Price Index ($CPI$) and Personal Consumption Expenditure implicit Deflator ($PCED$). The four forecasted measures of real activity are Personal Income ($PI$), Industrial Production ($IP$) index, Unemployment Rate ($UR$), and EMPloyees on non-farm Payrolls ($EMP$). Lastly, we consider forecasts for 3 month Treasury Bills as a measure of the short-term rate ($TBILL$) and 10 year Treasury Bonds as a measure of long-term rate ($TBOND$).

The series of interest are non-stationary and depending on their nature some transformations are adopted prior to forecasting. In particular, we distinguish among three categories:

- Prices: we forecast the $h$-months changes of yearly inflation. For instance, we forecast $(\pi_{t+h}^{CPI} - \pi_t^{CPI})$ for the consumer price index where $\pi_t^{CPI} = (\log(CPI_t) - \log(CPI_{t-12})) \times 100$.
- Industrial production, employees on non-farm payrolls and personal income: we forecast the $h$-months ahead annualized growth rate. For example we forecast $(1200/h) \times (\log(IP_{t+h}) - \log(IP_t))$ for the industrial production.
- Unemployment and interest rates: we forecast the $h$-months ahead changes. For instance we forecast $(UR_{t+h} - UR_t)$ for the unemployment rate.

Turning to the forecasting models, we consider the following specifications:

1. A $Naive$ forecast model in which forecasts of each (transformed) variable are simply a constant. This corresponds to a Random Walk ($RW$) model with drift for (i) the (log of) industrial production, personal income and employment and (ii) the rates of annual prices inflation, unemployment and interest rates. We will use interchangeably $Naive$ and $RW$.

2. Univariate forecasts ($AR$), where the forecasts are based exclusively on the own past values of the variable of interest.

3. Factor augmented $AR$ forecast ($FAAR$), in which the univariate models are augmented with common factors extracted from the whole panel of series.

4. Pooling of bivariate forecasts ($POOL$): for each variable the forecast is defined as the average of 130 forecasts obtained by augmenting the $AR$ model with each of the remaining 130 variables in the data set.

Pseudo out-of-sample forecasts are calculated for each variable and method over the horizons $h = 1, 3, 6$, and 12 months. The pseudo out-of-sample forecasting period begins in January 1970 and ends in December 2003. Forecasts constructed at date $T$ are based on models that are estimated using observations dated $T$ and earlier. We focus on rolling samples using, at each point in time, observations over the most recent 10 years.[2]

Rolling window estimators are attractive, in our context, for two reasons. First, they are better suited that recursive samples to investigate time variation in predictability. Second, large and persistent changes in the parameters of the models, like those associated with the Great Moderation, may result in less accurate estimates for the recursive samples.[3]

The measure for forecast evaluation is the Mean Square Forecast Error:

$$MSFE_{t_0}^{t_1}(i,h,m) = \frac{1}{t_1 - t_0 + 1} \sum_{t=t_0}^{t_1} \left( \hat{Y}_{i,t+h|T}^{h}(m) - Y_{t+h}^{h} \right)^2$$

where $1970:1 \leq t_0 \leq t_1 < 2003:12-h$. This is the average squared error between time $T_0$ and $T_1$, for variable $i$, at horizon $h$, using model $m$.

Following Granger and Newbold (1986), predictability is defined as the ratio between the MSFE of the forecasting model and a naive model:

$$PRED_{t_0}^{t_1}(i,h,m) = \frac{MSFE_{t_0}^{t_1}(i,h,m)}{MSFE_{t_0}^{t_1}(i,h,Naive)}$$

It is worth emphasizing that predictability conveys information on the *conditional variance* of the forecasting model and, therefore, provides additional information relative to the MSFE of a naive random walk model.

[2]Results are robust to alternative window width selections. In D'Agostino, Giannone and Surico (2006, Appendix C), we also report in the results for the recursive forecasts. In this case, the estimation period begins always in 1959M1.

[3]Rolling window estimators have the further advantage that they preserve the effect of estimation uncertainty on forecast performance. In contrast, estimation uncertainty vanishes asymptotically for expanding window methods such as recursive estimation schemes (see Giacomini and White, 2006).

# 3 Full-sample results

Our analysis begins with the full-sample evidence in Table 1. We report the predictability of four forecasting models, namely an AutoRegressive (AR) process, a Factor Augmented AutoRegressive (FAAR) forecast and a POOL of bivariate specifications. The naive, random walk, model is chosen as benchmark. The methods are displayed in blocks of rows. The first three columns refer to inflation, the central panel reports results for four measures of real activity while the last two columns are interest rates. Asterisks indicate a rejection of the test of equal predictive accuracy between each model and the random walk.[4]

Table 1: *Relative Mean Square Forecast Errors - Full Period*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ***Random Walk (absolute values)*** | | | | | | | | | |
| hor(m) | PPI | CPI | PCED | PI | IP | UR | EMP | TBILL | TBOND |
| 1 | 0.45 | 0.11 | 0.06 | 45.58 | 75.84 | 0.03 | 9.45 | 0.31 | 0.11 |
| 3 | 1.83 | 0.59 | 0.32 | 13.93 | 46.23 | 0.14 | 7.25 | 1.29 | 0.47 |
| 6 | 4.40 | 1.63 | 0.94 | 7.72 | 35.04 | 0.45 | 6.66 | 2.50 | 0.99 |
| 12 | 11.87 | 5.02 | 2.90 | 5.03 | 25.30 | 1.38 | 5.75 | 4.74 | 2.20 |
| ***Method AR (relative to RW)*** | | | | | | | | | |
| hor(m) | PPI | CPI | PCED | PI | IP | UR | EMP | TBILL | TBOND |
| 1 | 0.96 | $0.83^{***}$ | $0.83^{***}$ | 1.22 | $0.86^{*}$ | 0.91 | $0.60^{***}$ | 0.98 | 0.92 |
| 3 | 1.03 | $0.88^{*}$ | $0.82^{**}$ | 1.09 | 0.86 | $0.81^{*}$ | $0.53^{***}$ | 1.10 | 1.10 |
| 6 | 1.00 | 0.84 | 0.82 | 1.08 | 0.94 | 0.88 | $0.61^{***}$ | 1.05 | 1.05 |
| 12 | 1.05 | 0.93 | 1.00 | 1.01 | 0.95 | 0.97 | $0.75^{***}$ | 1.20 | 1.03 |
| ***Method FAAR (relative to RW)*** | | | | | | | | | |
| hor(m) | PPI | CPI | PCED | PI | IP | UR | EMP | TBILL | TBOND |
| 1 | 0.94 | $0.76^{***}$ | $0.78^{***}$ | 1.15 | $0.74^{***}$ | $0.72^{***}$ | $0.50^{***}$ | 0.93 | 0.95 |
| 3 | 0.91 | $0.71^{***}$ | $0.77^{**}$ | 0.93 | $0.64^{**}$ | $0.58^{***}$ | $0.39^{***}$ | 1.06 | 1.19 |
| 6 | 0.84 | $0.60^{***}$ | 0.75 | 0.90 | $0.63^{*}$ | $0.55^{***}$ | $0.43^{***}$ | 0.95 | 1.17 |
| 12 | 0.84 | $0.60^{*}$ | 0.83 | 0.94 | 0.63 | $0.64^{*}$ | $0.56^{***}$ | 1.05 | 1.26 |
| ***Method POOL (relative to RW)*** | | | | | | | | | |
| hor(m) | PPI | CPI | PCED | PI | IP | UR | EMP | TBILL | TBOND |
| 1 | 0.94 | $0.80^{***}$ | $0.80^{***}$ | 1.18 | $0.80^{**}$ | $0.83^{***}$ | $0.56^{***}$ | 0.94 | 0.91 |
| 3 | 0.96 | $0.81^{***}$ | $0.78^{**}$ | 1.02 | $0.76^{**}$ | $0.73^{**}$ | $0.47^{***}$ | 1.08 | 1.12 |
| 6 | 0.92 | $0.72^{**}$ | $0.76^{*}$ | 1.00 | $0.80^{*}$ | $0.76^{*}$ | $0.54^{***}$ | 0.99 | 1.07 |
| 12 | 0.92 | $0.73^{*}$ | 0.85 | $0.93^{**}$ | $0.78^{**}$ | $0.84^{***}$ | 0.65 | 1.12 | 1.07 |

Notes: Asterisks denote model forecasts that are statistically more accurate than the Naive at 1% ($***$), 5% ($**$) and 10% ($*$) significance levels.

For all prices and most real activity indicators, the forecasts based on large information are significantly more accurate than the Naive forecasts, with the factor augmented model producing the most accurate predictions. Univariate autoregressive forecasts significantly improve on the naive models for *EMP* at all horizons and for *CPI* and *PCED* at one and three month horizons only. As for interest rates, no forecasting model performs significantly better than the naive benchmark.

The evidence in Table 1 is consistent with the results in Stock and Watson (2005) and strongly supports the view that, in most situations, the non-benchmark models have a significant forecasting advantage relative to the naive models. This is the case for all predicted series with the exception of the short-term and long-term interest rates.

It is worth to emphasize that this kind of evaluations have been typically used in the literature as a model selection device for identifying the best forecasting method(s) in a pool

[4] Our inference is based on the regression: $\left(z_{ht} - \hat{z}_{ht}^{m}\right)^2 - \left(z_{ht} - \hat{z}_{ht}^{Naive}\right)^2 = c + u_{ht}$ where $z$ is the variable to be forecasted at horizon $h$ using *model-m*. The estimate of $c$ is simply the difference between *model-m* and a *Naive* model MSFEs, and the standard error is corrected for heteroskedasticity and serial correlation over $h-1$ months (see Romer and Romer, 2000). This testing procedure falls in the Diebold-Mariano-West framework, and Giacomini and White (2005, Section 3.2, see in particular Comment 4) show that by using rolling window estimators, as we do here, the limiting behavior of this type of tests is standard, and therefore standard asymptotic theory can be used for inference on the difference in predictive accuracy.

of alternative candidates. We show in the next section, however, that these findings are largely driven by the 1970s and the beginning of the 1980s when the majority of macroeconomic series were highly persistent and volatile. This observation appears to limit the benefit of performance evaluations over long sample periods that may be subject to parameter instability.

# 4 Forecast performance over sub-samples

In this section, we present evidence that the great moderation has been associated with a generalized decline in the predictability of several measures of inflation and real activity. Results for interest rates are also presented. We want to focus on the connection between changes in volatility and changes in predictability. This motivates a sample split around the mid-80s when most empirical contributions have dated the beginning of the Great Moderation (see for instance Kim and Nelson, 1999, and McConnell and Perez-Quiros, 2000, among many others). Interestingly, in a more recent study, Rossi and Sekhposyan (2008) search for break dates in relative forecast accuracy and find that the predictability of several US macroeconomic series declined around the mid-80s.

To assist the reader in evaluating the importance of the historical decline in predictability, we compute for each model the percentage change in the relative MSFEs between Period I, 1971-1984, and Period II, 1985-2002. Period II covers the great moderation sample. For each series and horizon, Tables 2 to 4 report the *average* percentage change among models. The statistics 'change' is defined in Appendix A.

## 4.1 Inflation

In Table 2, we report the results for all models. Moving from Period I to Period II, we note a great moderation in the forecast uncertainty associated with the naive model, as captured by the decline in the *absolute* values of the MSFE of the $RW$. On the other hand, the $AR$, $FAAR$ and $POOL$ models are associated with a remarkable decline in predictability with percentage changes, reported in the last column, of 40% magnitude on average. The largest changes are associated with six and twelve month horizons, especially for *CPI*.

In order to gauge the statistical significance of the historical changes in predictability using *rolling samples*, Table 2 reports asterisks whenever the forecast of a model is more accurate than the naive. At glance, the asterisks dominate the left part of Table 2. As for $CPI$ and $PCED$, the $AR$, $FAAR$ and $POOL$ methods significantly outperform the $RW$ before 1985. Furthermore, in line with Atkenson and Ohanian (2001), multivariate models appear to retain a forecasting advantage upon univariate models during the earlier period, especially at long horizons.

The finding of equal predictive accuracy during the last two decades is not specific to the best forecasting model, rather it appears a common feature of all methods. This observation leads to a new interpretation of the results in Atkenson and Ohanian (2001), Stock and Watson (2007) and D'Agostino and Giannone (2005) about the deterioration of the inflation forecasts on the basis of Phillips curve models and $FAAR$.

## 4.2 Real activity

We now turn the attention to the real side of the economy and investigate the properties of the forecasts of Personal Income ($PI$), Industrial Production ($IP$), Unemployment Rate ($UR$) and EMPloyees nonfarm payrolls ($EMP$). Table 3 reports the results.

Table 2: *Relative MSFEs across Sub-Periods - Inflation*

PERIOD I: sub-sample 1971:1 - 1984:12 PERIOD II: sub-sample 1985:1 - 2002:12 CHANGE

***Series: Producer Price Index***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.55 | 1.03 | 1.01 | 0.99 | 1 | 0.37 | $0.89^{*}$ | $0.87^{*}$ | $0.88^{***}$ | 7% |
| 3 | 2.23 | 1.05 | 0.85 | 0.94 | 3 | 1.51 | 1.01 | 0.98 | $0.99^{**}$ | 20% |
| 6 | 5.79 | 0.95 | 0.67 | $0.82^{**}$ | 6 | 3.31 | 1.08 | 1.08 | 1.07 | 34% |
| 12 | 17.95 | 1.02 | 0.65 | 0.84 | 12 | 7.12 | 1.13 | 1.20 | 1.09 | 33% |

***Series: Consumer Price Index***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.17 | $0.83^{***}$ | $0.75^{***}$ | $0.78^{***}$ | 1 | 0.07 | $0.85^{*}$ | $0.77^{**}$ | $0.83^{***}$ | 5% |
| 3 | 0.94 | $0.84^{*}$ | $0.61^{***}$ | $0.74^{***}$ | 3 | 0.31 | 0.99 | 0.93 | $0.96^{**}$ | 38% |
| 6 | 2.85 | $0.78^{*}$ | $0.46^{***}$ | $0.65^{***}$ | 6 | 0.68 | 1.04 | 1.05 | $0.98^{*}$ | 83% |
| 12 | 9.43 | 0.87 | $0.44^{***}$ | $0.64^{**}$ | 12 | 1.57 | 1.22 | 1.32 | 1.16 | 118% |

***Series: Personal Consumption Expenditure Deflator***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.08 | $0.73^{***}$ | $0.71^{***}$ | $0.71^{***}$ | 1 | 0.05 | 0.96 | $0.88^{**}$ | $0.93^{***}$ | 9% |
| 3 | 0.50 | $0.72^{***}$ | $0.67^{**}$ | $0.68^{***}$ | 3 | 0.18 | 1.04 | 0.98 | 1.01 | 29% |
| 6 | 1.63 | $0.72^{**}$ | $0.66^{*}$ | $0.66^{**}$ | 6 | 0.40 | 1.13 | 1.05 | 1.08 | 48% |
| 12 | 5.52 | 0.92 | 0.75 | 0.77 | 12 | 0.85 | 1.37 | 1.27 | 1.27 | 59% |

Notes: The column 'change' reads the percentage historical decline in predictability averaged across methods (excluding Naive). Asterisks denote model forecasts that are statistically more accurate than the Naive at 1% ($***$), 5% ($**$) and 10% ($*$) significance levels.

Table 3: *Relative MSFEs across Sub-Periods - Real Activity*

PERIOD I: sub-sample 1971:1 - 1984:12 PERIOD II: sub-sample 1985:1 - 2002:12 CHANGE

***Series: Real Personal Income***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 38.54 | 1.02 | 0.95 | 0.98 | 1 | 51.09 | 1.33 | 1.27 | 1.30 | 21% |
| 3 | 17.15 | 1.01 | 0.86 | 0.94 | 3 | 11.41 | 1.19 | 1.01 | 1.12 | 14% |
| 6 | 10.41 | 1.05 | 0.83 | 0.96 | 6 | 5.62 | 1.12 | 1.01 | 1.05 | 2% |
| 12 | 6.92 | 0.97 | 0.84 | $0.87^{*}$ | 12 | 3.55 | 1.07 | 1.09 | 1.02 | 3% |

***Series: Industrial Production***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 124.01 | $0.81^{*}$ | $0.65^{***}$ | $0.75^{**}$ | 1 | 38.14 | 0.97 | 0.95 | 0.92 | 14% |
| 3 | 81.48 | 0.85 | $0.55^{**}$ | $0.73^{**}$ | 3 | 18.64 | 0.92 | 0.98 | 0.88 | 16% |
| 6 | 61.42 | 0.94 | $0.49^{*}$ | $0.76^{*}$ | 6 | 14.41 | 0.97 | 1.11 | 0.95 | 34% |
| 12 | 43.24 | 0.95 | $0.43^{**}$ | $0.72^{**}$ | 12 | 11.27 | 0.98 | 1.22 | 0.97 | 62% |

***Series: Unemployment Rate***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.05 | 0.86 | $0.63^{***}$ | $0.78^{**}$ | 1 | 0.02 | 0.99 | $0.88^{*}$ | $0.94^{***}$ | 21% |
| 3 | 0.25 | 0.79 | $0.52^{***}$ | $0.69^{**}$ | 3 | 0.06 | 0.91 | $0.79^{*}$ | $0.84^{**}$ | 18% |
| 6 | 0.80 | 0.88 | $0.49^{***}$ | 0.75 | 6 | 0.17 | 0.85 | 0.75 | $0.80^{*}$ | 22% |
| 12 | 2.42 | 0.99 | $0.56^{**}$ | $0.82^{**}$ | 12 | 0.56 | 0.93 | 0.90 | 0.89 | 41% |

***Series: Employees on Nonfarm Payrolls***

| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 16.37 | $0.65^{***}$ | $0.51^{***}$ | $0.60^{***}$ | 1 | 4.04 | $0.42^{***}$ | $0.45^{***}$ | $0.40^{***}$ | 4% |
| 3 | 12.39 | $0.60^{**}$ | $0.41^{***}$ | $0.53^{***}$ | 3 | 3.23 | $0.31^{***}$ | $0.34^{***}$ | $0.29^{***}$ | -1% |
| 6 | 11.16 | $0.70^{**}$ | $0.42^{***}$ | $0.60^{**}$ | 6 | 3.14 | $0.37^{**}$ | $0.44^{*}$ | $0.36^{*}$ | -3% |
| 12 | 9.21 | $0.82^{***}$ | $0.49^{***}$ | $0.69^{***}$ | 12 | 3.05 | $0.58^{**}$ | 0.72 | 0.56 | 8% |

Notes: see Table 2.

The reduction in forecast uncertainty associated with the Great Moderation is apparent in the decline of the absolute MSFEs of the $RW$ for all variables and horizons, with the exception of real personal income one-month ahead. As for the relative performance of sophisticated models, the $FAAR$ is the best predictive model in Period I. The significant forecasting advantage over the earlier sample, however, is sizably reduced over Period II. Furthermore, the historical changes in the last column are sizable, around 20% on average, and the predictions of $FAAR$

and *POOL* are always more accurate than the naive model before 1985.

In analogy to the results for inflation, the left panel of Table 3, which refers to the earlier subsample, is dominated by asterisks. In contrast to Table 2, univariate *AR* specifications for *PI*, *IP* and *UR* poorly perform even before 1985 and the null hypothesis of equal predictive accuracy relative to the *RW* is not rejected over both samples. On the other hand, the *FAAR* and *POOL* methods produce significantly more accurate forecasts during Period I.

The relative MSFEs of *AR* over the two subsamples confirm the result in Stock and Watson (2003a) of little change in the structure of *univariate* models for real activity. The relative MSFEs of *FAAR* and *POOL*, however, suggest that important changes have occurred in the relationship between output and other macroeconomic variables.[5]

It is interesting to notice that the decline in predictability does not seem to extend to the labor market, especially at short horizons. The forecasts of the employees on nonfarm payrolls are associated with the smallest percentage changes across subsamples. Furthermore, the relative MSFEs of most models are statistically different from one in both Periods.

The evidence presented in this section corroborates the view that the decline in predictability is intrinsic to the post-1985 data rather than specific to a particular forecasting model.[6]

## 4.3 Interest rates

The behaviour of the interest rate forecasts in Table 4 contrasts with the behaviour of all other variables across sub-samples, especially at the very short horizon. In particular, the average *increases* in the predictability of the short-term rate are 10% and 5% for $h = 1$ and 3, being among the very few percentage changes with a negative sign. The *POOL* forecasts are characterized by the most pronounced historical improvement and become more accurate than the RW in the most recent period. At the longer horizons of six and twelve months, however, the relative MSFEs remain above one.

Table 4: *Relative MSFEs across Sub-Periods - Interest Rates*

| PERIOD I: sub-sample 1971:1 - 1984:12 | | | | | PERIOD II: sub-sample 1985:1 - 2002:12 | | | | | CHANGE |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Series: 3 Months Treasury Bills*** | | | | | | | | | | |
| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
| 1 | 0.64 | 1.00 | 0.94 | 0.95 | 1 | 0.05 | 0.84 | 0.87 | 0.81*** | -10% |
| 3 | 2.59 | 1.12 | 1.05 | 1.10 | 3 | 0.27 | 0.98 | 1.16 | 0.94** | -5% |
| 6 | 4.63 | 1.06 | 0.88 | 0.98 | 6 | 0.83 | 1.03 | 1.25 | 1.01 | 11% |
| 12 | 7.63 | 1.27 | 0.93 | 1.14 | 12 | 2.47 | 1.04 | 1.34 | 1.06 | 8% |
| ***Series: 10 Years Treasury Bonds*** | | | | | | | | | | |
| hor | RW | AR | FAAR | POOL | hor | RW | AR | FAAR | POOL | Average |
| 1 | 0.17 | 0.95 | 0.96 | 0.94 | 1 | 0.07 | 0.88** | 0.92 | 0.87*** | -9% |
| 3 | 0.68 | 1.17 | 1.21 | 1.18 | 3 | 0.31 | 1.00 | 1.15 | 1.02 | -11% |
| 6 | 1.28 | 1.07 | 1.12 | 1.09 | 6 | 0.77 | 1.02 | 1.23 | 1.05 | 3% |
| 12 | 2.57 | 1.04 | 1.12 | 1.06 | 12 | 1.91 | 1.01 | 1.42 | 1.09 | 7% |

Notes: see Table 2.

It is worth emphasizing that ending the earlier sub-sample in 1979:10, which corresponds to the beginning of Volcker's experiment of non-borrowed reserve targeting, does not overturn the result on interest rate unpredictability at short horizons.[7] Results are available upon request.

[5] Giannone, Lenza and Reichlin (2007) reach a similar conclusion on the change of the propagation mechanism using a VAR with nineteen variables. We return to this in Section 6.

[6] De Mol, Giannone and Reichlin (2007) report a similar finding using Bayesian forecasts methods.

[7] Excluding the period of Volcker's experiment from period I improves the forecast ability one year ahead.

The absolute MSFEs of the *RW* fall for the long-term interest rate, though the historical decline is less pronounced than for the short-term rate. The other methods produce significantly more accurate one-month ahead forecasts in Period II, consistently with the results on the 3 months treasury bills. At longer horizons, however, the performance of all forecasting models is very close to the performance of RW. The latter finding holds over both sub-samples and thus extends the results of interest rate unpredictability at long horizons that Rudebusch (2002) reports for Greenspan's tenure only.

In summary, during Period II the *AR*, *FAAR* and *POOL* methods produce more accurate forecasts than the *RW* at the very short horizon of *one month.* An interesting interpretation of this result is that a stronger policy activism and a better communication strategy have enriched the information content of the systematic component of monetary policy during the last two decades. Indeed, the St. Louis Fed President William Poole (2005) mentions the increase in transparency, and the consequent increase in predictability of monetary policy among the four identifying characteristics of the Greenspan era and argues that "*[..] improved predictability of policy has had much to do with improved effectiveness of policy*". Empirical support for the improved effectiveness of U.S. monetary policy can be found in Boivin and Giannoni (2006).

# 5 Evidence from institutional forecasters

Taking the results of the previous section at face value, we might conclude that inflation and real activity have became less predictable since 1985. While this claim appears valid across several statistical methods, it is less clear the extent to which it applies to larger, possibly non-linear models such as those employed by Central Banks and private forecasters. The forecasts produced by policy institutions are likely to involve some important elements of judgement that can improve predictive accuracy relative to more mechanical methods.

## 5.1 The Federal Reserve and the professional forecasters

We consider the predictions for output and its deflator from two large forecasters representing the private sector and the policy institutions. The source for the commercial providers is the Survey of Professional Forecasts (SPF). The survey was introduced by the American Statistical Association and the National Bureau of Economic Research and is currently maintained by the Philadelphia Fed. The SPF refers to quarterly measures and is conducted in the middle of the second month of each quarter. We consider the median of the individual forecasts.[8]

As far as institutional forecasts are concerned, we consider the forecasts of the Greenbook. These forecasts are prepared by the Board of Governors at the Federal Reserve for the meetings of the Federal Open Market Committee (FOCM), which takes place roughly every six weeks. The predicted series are quarterly inflation and output. The Greenbook forecast are made publicly available with a five-year delay, thereby implying that our sample ends in 1999. For comparability with the timing of the SPF forecasts, we select meetings that are closer to the middle of each quarter (i.e. four meeting out of eight).

We consider four forecast horizons $h_q$ ranging from 1 to 4 *quarters.* The one step ahead figures correspond to the predictions for the quarter in which the forecasts are made. For each $h_q$-steps ahead we consider the $h_q$-quarter growth rate of output and the $h_q$-quarter change in annual inflation based on the output implicit price deflator. The measure of output is Gross National Product (GNP) until 1991 and Gross Domestic Product (GDP) from 1992 onwards. The evaluation sample begins in 1975, as prior to this date the Greenbook forecasts were not

[8]The data are available from the website of the Federal Reserve Bank of Philadelphia. http://www.phil.frb.org/econ/spf/spfmed.html; http://www.phil.frb.org/econ/forecast/croushoresdatasets.html; http://www.phil.frb.org/econ/forecast/reaindex.html.

always available up to the fourth quarter horizon. For the sake of comparability, we select 1975 as starting point also for the SPF forecasts, although the latter are available for a longer time period. Data are continuously revised and thus for each quarter several measures of inflation and output are available. Following Romer and Romer (2000), we consider the figures published after the next two subsequent quarters.

Finally, the Naive forecasts are computed as the sample average of the $h_q$-quarter growth rate of output and the $h_q$-quarter change of annual inflation based on the output implicit price deflator. In line with the forecasts of the statistical methods, the parameters of the Naive forecasts are computed using observations over the most recent 10 years. We use real-time data as available to the Fed when the GB forecasts were actually produced.

## 5.2 The decline of predictability

We turn now to the evaluation of the forecasts produced by the Federal Reserve and the SPF over inflation and real activity relative to a naive random walk model. Our goal is to assess the robustness of the historical decline in predictability by asking whether this finding is independent from the model at hand. Results for inflation and output are presented in Table 5 and Table 6. The statistics refer to three periods: full sample, pre-1985 and post-1985 periods.

Table 5: *Relative MSFEs of Institutional Forecasters - Inflation*

***FULL SAMPLE: 1975:1 1999:4***

| hor(q) | RW | Fed's Green Book(GB)/RW | Survey of Professional Forecasters(SPF)/RW |
|---|---|---|---|
| 1 | 0.26 | $0.35^{***}$ | $0.37^{***}$ |
| 2 | 0.79 | $0.30^{**}$ | $0.36^{**}$ |
| 3 | 1.57 | $0.29^{*}$ | 0.37 |
| 4 | 2.51 | 0.32 | 0.46 |

***PERIOD I: sub-sample 1975:1 - 1984:4***

| hor(q) | RW | Fed's Green Book(GB)/RW | Survey of Professional Forecasters(SPF)/RW |
|---|---|---|---|
| 1 | 0.54 | $0.30^{***}$ | $0.27^{***}$ |
| 2 | 1.72 | $0.21^{**}$ | $0.24^{**}$ |
| 3 | 3.51 | $0.21^{**}$ | $0.25^{*}$ |
| 4 | 5.69 | $0.23^{*}$ | $0.32^{*}$ |

***PERIOD II: sub-sample 1985:1 - 1999:4***

| hor(q) | RW | Fed's Green Book(GB)/RW | Survey of Professional Forecasters(SPF)/RW |
|---|---|---|---|
| 1 | 0.08 | $0.58^{**}$ | 0.82 |
| 2 | 0.17 | 0.93 | 1.15 |
| 3 | 0.28 | 0.97 | 1.39 |
| 4 | 0.39 | 1.18 | 1.82 |

Notes: Asterisks denote rejection of the null hypothesis of equal predictive accuracy between each model and the RW at 1% ($***$), 5% ($**$) and 10% ($*$) significance levels.

The top panel of Table 5 presents the finding for the full sample. For inflation, the Greenbook and the SPF forecasts are far more accurate than a naive model, being associated with significantly lower MSFEs at all horizons. The results of Period I in the middle panel are very similar to the full-sample results whereas for the post-1985 period the statistics in the bottom panel paint a different picture. In particular, the relative MSFEs of Period II are very close to *one* for most horizons, and the null hypothesis of equal predictive accuracy between the naive model and the other forecasts is not rejected in all cases but $h_q = 1$ for the Greenbook.

The results for real output are displayed in Table 6 and they bear out the evidence on inflation. In particular, the forecasts of the Greenbook and the SPF are significantly more accurate than the RW over the full-sample and the earlier period. After 1985, however, the statistics in the last row are associated with relative MSFEs close to *one*, thereby revealing

Table 6: *Relative MSFEs of Institutional Forecasters - Output*

| hor(q) | RW | Fed's Green Book(GB)/RW | Survey of Professional Forecasters(SPF)/RW |
|---|---|---|---|
| ***FULL SAMPLE: 1975:1 1999:4*** | | | |
| 1 | 12.59 | $0.44^{**}$ | $0.51^{**}$ |
| 2 | 9.11 | $0.49^{**}$ | $0.46^{**}$ |
| 3 | 7.45 | $0.48^{**}$ | $0.50^{***}$ |
| 4 | 6.49 | $0.51^{**}$ | $0.51^{***}$ |
| ***PERIOD I: sub-sample 1975:1 - 1984:4*** | | | |
| 1 | 25.82 | $0.37^{**}$ | $0.45^{**}$ |
| 2 | 19.01 | $0.44^{**}$ | $0.41^{**}$ |
| 3 | 15.39 | $0.40^{***}$ | $0.45^{***}$ |
| 4 | 13.18 | $0.42^{***}$ | $0.46^{***}$ |
| ***PERIOD II: sub-sample 1985:1 - 1999:4*** | | | |
| 1 | 3.77 | 0.73 | 0.77 |
| 2 | 2.51 | 0.77 | 0.70 |
| 3 | 2.15 | 0.85 | 0.73 |
| 4 | 2.03 | 0.89 | 0.74 |

Notes: see Table 5.

that more sophisticated forecasts for output are not immune to the generalized decline in predictability.[9]

These findings complement the statistics of the previous section and disclose two new results. First, in analogy to the statistical models, the performance of both the Greenbook and SPF over the full-sample are mainly driven by the time period before 1985. Second, the Greenbook and the SPF forecasts are characterized by a significant decline in the predictability such that the advantage of the 1970s and the first half of the 1980s relative to a naive model has virtually vanished during the last two decades.

Unlike the statistical models, however, during the later subsample the Greenbook retains some advantage over naive forecasts at the very short horizon of one quarter.[10] An explanation for this result is that the models employed by the Fed are flexible enough to use the high frequency information available within a quarter for predicting the current values of other series. This feature makes large models particularly helpful for conjunctural analysis.[11]

# 6 What drives the decline in predictability?

In this section, we explore whether the decline in macroeconomic predictability documented in this paper is mainly driven by changes in the estimated coefficients or by changes in the estimated error variances of our statistical models. There are many ways in which it is possible to construct a sensible counterfactual experiment, and with so many forecasting models and variables as in this paper the choice is not obvious.

A simple experiment that may shed some lights on the sources of the decline in predictability goes as follows. First, in each sub-sample we extract the common factors from the whole panel of series. For consistency with the analysis above, we select three factors. Second, for each period we fit a VAR on the estimated factors augmented with the variable we want to forecast. Third, we evaluate, out-of-sample, the forecasts of the VAR relative to the forecasts of the naive model using the RMSFE statistics. The results are reported in the first two rows of Table 7 and show that, in this experiment, the forecasts based on the VAR replicate the fall

[9]A similar result for SPF predictions on output growth can be found in Campbell (2007). The focus of that paper, however, is on reduced macroeconomic uncertainty rather than on the predictability of widely used forecasting models for inflation and real activity.

[10]See D'Agostino and Whelan (2007) for a detailed description of the evolution of the Fed advantage to forecast inflation and GDP growth before and after 1992.

[11]Giannone, Reichlin and Small (2007) formalize these procedures in a data-rich environment.

in predictability documented in section 4. Fourth, we super-impose the VAR coefficients (error variances) estimated over the simulated period I onto the VAR error variances (coefficients) estimated over the simulated period II and then compute, via bootstrap, the relative mean squared forecast errors implied, out-of-sample, by the counterfactual VAR. In the third (fourth) row of Table 7, we report the average RMSFEs over 2000 repetitions for VARs of order 4. The message of this section, however, does depend neither on the specific selection of horizons and order lags, nor on estimating the factors over sub-periods as opposed to the full-sample.

Table 7: *Relative Mean Square Forecast Errors - counterfactuals*

| coefficients | variances | PPI | CPI | PCED | PI | IP | UR | EMP | TBILL | TBOND |
|---|---|---|---|---|---|---|---|---|---|---|
| period I | period I | 0.78 | 0.42 | 0.47 | 0.43 | 0.83 | 1.14 | 0.60 | 0.45 | 0.69 |
| period II | period II | 1.06 | 1.03 | 0.68 | 0.67 | 1.00 | 1.02 | 0.88 | 0.97 | 1.05 |
| period I | period II | 0.75 | 0.49 | 0.53 | 0.39 | 0.60 | 1.08 | 0.43 | 0.41 | 0.55 |
| period II | period I | 0.76 | 0.90 | 0.80 | 0.57 | 0.97 | 1.03 | 0.91 | 0.65 | 0.72 |

If a change in the estimated coefficients was the main driver of the decline in predictability, then we would expect: (i) little change in the RMSFEs moving from the first to the third row, (ii) an *increase* in the RMSFEs moving from the first to the fourth row. If, on the other hand, a change in the estimated error variances was behind the fall in forecast ability, then we would expect: (i) little change in the RMSFEs moving from the second to the third row, (ii) a *decrease* in the RMSFEs moving from the second to the fourth row.

A comparison between the second and third rows of Table 7 suggests that granting period II with the coefficients from period I overturns the finding of a decline in predictability. From the first and fourth row, in contrast, it emerges that granting period I with the coefficients from period II is sufficient to generate a decline in predictability for most series.[12] In summary, the results of this section suggest that the change in the autoregressive coefficients has been a main driver of the decline in macroeconomic predictability.

# 7 Conclusions

In this paper, we have investigated the ability of some widely used econometric models, the Fed's Greenbook and the Survey of Professional Forecasters to predict several U.S. macroeconomic time series. A main result is that, moving from the pre- to the post-1985 period, there is a sizable and significant deterioration in the forecast accuracy of these methods *relative* to a naive random walk model. This finding is robust across forecast horizons and models, and applies also to the predictions of inflation and output made by the Fed. In particular, during the last two decades, more sophisticated methods such as those contributing to the Greenbook offer no higher predictive accuracy than do naive forecasts for all horizons but the first quarter.

It is worth to emphasize, however, that our findings should not be interpreted as suggestive that forecasting can be regarded as unimportant in modern policy making. The out of sample performance of a model *in real time* is in fact a far more complex evaluation than our *ex-post* analysis could capture. As long as there exists some positive probability that the current macroeconomic stability may come to an end, policy institutions like Central Banks will have strong incentives to devote resources to forecast inflation and output, because it is in those times that their comparative advantage emerges. Furthermore, within the current quarter, which is arguably the relevant horizon for conjunctural analysis, the Fed's Greenbook continues to maintain a forecasting advantage relative to less sophisticated models.

[12] The sub-samples used for the simulations behind Table 7 are of the same size of the sub-samples used on actual data. To ensure that our results do not depend on estimation uncertainty, we repeat the counterfactual exercises generating samples of 2000 observations. Results are robust to this modification.

## References


[1] Ahmed Shaghil , Andrew T. Levin and Beth Anne Wilson (2004) "Recent U.S. Macroeconomic Stability: Good Policies, Good Practices, or Good Luck?", *The Review of Economics and Statistics*, 86, pp. 824-832.

[2] Atkenson Andrew and Lee E. Ohanian (2001), "Are Phillips Curves Useful for Forecasting Inflation?", *Federal Reserve Bank of Minneapolis Quarterly Review*, 25 (1).

[3] Bates, J.M., Clive W.J. Granger (1969), "The Combination of Forecasts ", *Operations Research Quarterly*, 20, 451-68.

[4] Benati Luca, and Paolo Surico (2008), "Evolving U.S. Monetary Policy and the Decline of Inflation Predictability", *Journal of the European Economic Association*, forthcoming.

[5] Blanchard, Oliver J. and John A. Simon (2001), "The Long and Large Decline in U.S. Output Volatility", *Brookings Papers on Economic Activity,* 1, 135-164.

[6] Boivin, Jean and Serena Ng (2005), "Understanding and Comparing Factor-Based Forecasts", *International Journal of Central Banking* 1, pp. 117-151.

[7] Boivin, Jean and Marc Giannoni (2006), "Has Monetary Policy Become more effective?", *The Review of Economic and Statistics* 88, pp. 445-462.

[8] Campbell, Sean, D., (2007), "Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters", *Journal of Business and Economic Statistics*, 25, pp. 191-200.

[9] Cogley, Timothy W. and Thomas Sargent, (2002) "Evolving Post World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual* 16.

[10] D'Agostino, Antonello and Domenico Giannone, (2006), "Comparing Alternative Predictors Based on Large-Panel Factor Models", European Central Bank working paper No. 680.

[11] D'Agostino, Antonello, Domenico Giannone and Paolo Surico, (2006), "(Un)Predictability and Macroeconomic Stability", European Central Bank working paper No. 605.

[12] D'Agostino, Antonello and Karl Whelan, (2008), "Federal Reserve Information during the Great Moderation", *Journal of the European Economic Association*, forthcoming.

[13] De Mol, Christine, Domenico Giannone and Lucrezia Reichlin, 2007, "Forecasting using a large number of predictors: is Bayesian regression a valid alternative to principal components?", *Journal of Econometrics*, forthcoming.

[14] Diebold, Francis X., and Lutz Kilian, (2001), "Measuring Predictability: Theory and Macroeconomic Applications", *Journal of Applied Econometrics* 16, pp. 657-669.

[15] Diebold, Francis X., and R.S. Mariano, (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, pp. 253-265.

[16] Estrella, Arturo, (2005), "Why Does the Yield Curve Predict Output and Inflation", *The Economic Journal* 115, pp. 722-744.

[17] Forni Mario, Marc Hallin, Marco Lippi and Lucrezia Reichlin, (2002) "The Generalized Dynamic Factor Model: Identification and Estimation", *The Review of Economics and Statistics* 82, pp. 540-554.

[18] Giacomini, Raffaella, and Halbert White, (2006), "Tests of Conditional Predictive Ability", *Econometrica* 74, pp. 1545-1578.

[19] Giannone, Domenico, Lucrezia Reichlin and David Small, (2007), "Nowcasting GDP and Inflation: The Real Time Informational Content of Macroeconomic Data Releases", *Journal of Monetary Economics*, forthcoming.

[20] Giannone, Domenico, Michele Lenza and Lucrezia Reichlin, (2008), "Explaining the Great Moderation: it is not the shocks", *Journal of the European Economic Association*, forthcoming.

[21] Granger, Clive W.J. and Paul Newbold, (1986), "Forecasting Economic Time Series", 2nd ed., Academic Press: Orlando, Florida.

[22] Kim, Chang-Jin, and Charles R. Nelson (1999) "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle", *The Review of Economics and Statistics*, 81, pp. 608-616.

[23] McConnell, Margaret M, and Gabriel Perez-Quiros (2000), "Output Fluctuations in the United States: What has Changed Since the Early 1980's", *American Economic Review* 90, 1464 – 1476.

[24] Poole, William (2005), Remarks: Panel on "After Greenspan: Whither Fed Policy?" — Western Economic Association International Conference (WEAI), San Francisco, July 6, 2005.

[25] Romer, David, H. and Christina Romer (2000), "Federal Reserve Information and the Behavior of Interest Rates", *American Economic Review* 90, 429-457.

[26] Rossi, Barbara Tatevik Sekhposyan, (2008), "Has Models' Forecasting Performance for US Output Growth and Inflation Changed over Time, and When?", mimeo, Duke University.

[27] Rudebusch, Glenn, (2002), "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia", *Journal of Monetary Economics* 49, September 2002, pp. 1161-1187

[28] Stock, James and Mark Watson (2005), "An Empirical Comparison of Methods for Forecasting Using Many Predictors", mimeo, Harvard University.

[29] Stock, James and Mark Watson (2007), "Why Has Inflation Become Harder to Forecasts?", *Journal of Money, Credit and Banking* 39, pp. 3-34.

[30] Stock, James and Mark Watson (2003a), "Has the Business Cycle Changed? Evidence and Explanations", prepared for Federal Reserve Bank of Kansas City 2003 Jackson Hole Symposium.

[31] Stock, James and Mark Watson (2003b), "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economics Statistics* 20, 147-42.

[32] Stock, James and Mark Watson (2003c), "Forecasting Output and Inflation: the Role of Asset Prices" *Journal of Economic Literature XLI*, pp. 788-829.

[33] West, Kenneth, D. (1996), "Asymptotic Inference about Predictive Ability" *Econometrica* 64, pp. 1067-1084.

## Appendix A: the Forecasting Models

We are interested in predicting some variable $Y^h_{i,t+h}$ using a potentially large number of predictors, $X_{i,t}, i = 1, ..., n$. To this end, we consider the following forecasting models:

**Naive**

$$Y^h_{i,t+h} = \alpha_i^{h,Naive} + e_{i,t+h}^{h,Naive}$$

**Autoregressive**

$$Y^h_{i,t+h} = \alpha_i^{h,AR} + \gamma_i^{h,AR}(L)X_{i,t} + e_{t+h}^{h,AR}$$

**Augmented distributed lag**

$$Y^h_{i,t+h} = \alpha_i^{h,ADL_j} + \gamma_i^{h,ADL_j}(L)X_{i,t} + \delta_j^{h,ADL_j}(L)X_{j,t} + e_{t+h}^{h,ADL_j}, j = 1, ..., n, j \neq i$$

**r-factor model**

$$Y^h_{i,t+h} = \alpha_i^{h,FAAR} + \gamma_i^{h,FAAR}(L)X_{i,t} + \lambda_i^{h,FAAR}\hat{F}_t + e_{t+h}^{h,FAAR}$$

The series are transformed by taking logarithms and/or differences. In general, growth rates are used for real quantity variables, first differences are used for nominal interest rates, and first differences for yearly growth rates of the price series.

Following Stock and Watson (2005), we set the lag length of all filters to 4 and the number of factors to 3. Their choice is based on the results of several different selection methods on similar models and the same dataset used in this paper.

Table A shows the definition of $Y^h_{i,t+h}$ and $X_{i,t}$ in terms of the raw series $Z_{it}$ for each of the nine variables that are forecasted. The transformations were used for all predictors listed in Appendix B.

Table A: Forecasted Series

| *Series* | *Acronyms* | $Y^h_{t+h}$ | $X_t$ |
|---|---|---|---|
| Real Personal Income | PI | $\left(\frac{1200}{h}\right)\ln\left(\frac{Z_{t+h}}{Z_t}\right)$ | $\Delta\ln(Z_t)$ |
| Industrial Production | IP | $\left(\frac{1200}{h}\right)\ln\left(\frac{Z_{t+h}}{Z_t}\right)$ | $\Delta\ln(Z_t)$ |
| Unemployment Rate | UR | $Z_{t+h} - Z_t$ | $\Delta Z_t$ |
| Employment | EMP | $\left(\frac{1200}{h}\right)\ln\left(\frac{Z_{t+h}}{Z_t}\right)$ | $\Delta\ln(Z_t)$ |
| 3-Mth Tbill Rate | TBILL | $Z_{t+h} - Z_t$ | $\Delta Z_t$ |
| 10-Yr Tbond Rate | TBOND | $Z_{t+h} - Z_t$ | $\Delta Z_t$ |
| Producer Price Index | PPI | $100\times\ln\left(\frac{Z_{t+12+h}}{Z_{t+h}}\right) - 100\times\ln\left(\frac{Z_{t+12}}{Z_{t-12}}\right)$ | $\Delta\ln\left(\frac{Z_t}{Z_{t-12}}\right)$ |
| Consumer Price Index | CPI | $100\times\ln\left(\frac{Z_{t+12+h}}{Z_{t+h}}\right) - 100\times\ln\left(\frac{Z_{t+12}}{Z_{t-12}}\right)$ | $\Delta\ln\left(\frac{Z_t}{Z_{t-12}}\right)$ |
| PCE Deflator | PCED | $100\times\ln\left(\frac{Z_{t+12+h}}{Z_{t+h}}\right) - 100\times\ln\left(\frac{Z_{t+12}}{Z_{t-12}}\right)$ | $\Delta\ln\left(\frac{Z_t}{Z_{t-12}}\right)$ |

Notes: This table lists the nine forecasted series. The first column gives the description of the series, the second lists the abbreviation used in the results tables, the next two columns shows the transformations that define the variable forecast, $Y_{t+h}$ and the predictors $X$.

Given a sample $t = T_{0T}, ..., T$, we estimate the common factors $\hat{F}_t$ by mean of the first $r$ sample principal components of $W_t = (W_{1t}, ..., W_{nt})', t = T_{0,T}, ..., T$, where $W_{it} = \frac{X_{it} - \hat{\mu}_i}{\hat{\sigma}_i}$, and $\hat{\mu}_i$ and $\hat{\sigma}_i$ are the sample mean and standard deviation respectively. Specifically, $\hat{F}_t = \hat{\mathcal{V}}'W_t$, where $\hat{\mathcal{V}}$ is the $n \times r$ matrix of eigenvectors associated with the first $r$ eigenvalues of $S = \frac{1}{T - T_{0T} + 1}\sum_{t=T_{0T}}^{T} W_t W_t'$.

The parameters of the each model can be thus computed by Ordinary Least Square. We obtain the following forecasts:

$$\hat{Y}^h_{i,T+h|T}(Naive) = \hat{\alpha}^{h,Naive}_i$$

$$\hat{Y}^h_{i,t+h|T}(AR) = \hat{\alpha}^{h,AR}_i + \hat{\gamma}^{h,AR}_i(L)X_{i,T}$$

$$\hat{Y}^h_{i,T+h|T}(ADL_j) = \hat{\alpha}^{h,ADL_j}_i + \hat{\gamma}^{h,ADL_j}_i(L)X_{i,T} + \hat{\delta}^{h,ADL_j}_j(L)X_{j,T}, j = 1, ..., n, j \neq i$$

$$\hat{Y}^h_{i,t+h|T}(FAAR) = \hat{\alpha}^{h,FAAR}_i + \hat{\gamma}^{h,FAAR}_i(L)X_{i,T} + \hat{\lambda}^{h,FAAR}_i\hat{F}_T$$

Pooled forecasts from different ADL models are computed as:

$$\hat{Y}^h_{i,t+h|T}(POOL) = \sum_{j\neq i}\hat{Y}^h_{i,t+h|T}(ADL_j)$$

The percentage decline in the relative MSFE of the *i-th* predicted series is averaged across models excluding the RW, and is computed as:

$$CHANGE\,(i,h) = 100\left[\frac{\sum_{m=1}^{M}\left(\frac{PRED^{II}(i,h,m)-PRED^{I}(i,h,m)}{PRED^{I}(i,h,m)}\right)}{M}\right]$$

with $m = AR$, $FAAR$ and $POOL$, the number of models $M = 3$ and $h = 1$, 3, 6 and 12.

## Appendix B: the Data Set

**Table B: Data Transformation**

| | Definition | Transformation |
|---|---|---|
| 1 | $X_{it} = Z_{it}$ | no transformation |
| 2 | $X_{it} = \Delta Z_{it}$ | monthly difference |
| 4 | $X_{it} = \ln Z_{it}$ | log |
| 5 | $X_{it} = \Delta \ln Z_{it} \times 100$ | monthly growth rate |
| 6 | $X_{it} = \Delta \ln \frac{Z_{it}}{Z_{it-12}} \times 100$ | monthly difference of yearly growth rate |

| Code | Description | Transf. |
|---|---|---|
| A0M051 | Personal income less transfer payments (AR, bil. chain 2000 $) | 5 |
| A0M224R | Real Consumption (AC) A0m224/gmdc | 5 |
| A0M057 | Manufacturing and trade sales (mil. Chain 1996 $) | 5 |
| A0M059 | Sales of retail stores (mil. Chain 2000 $) | 5 |
| IPS10 | INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX | 5 |
| IPS11 | INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL | 5 |
| IPS299 | INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS | 5 |
| IPS12 | INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS | 5 |
| IPS13 | INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS | 5 |
| IPS18 | INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS | 5 |
| IPS25 | INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT | 5 |
| IPS32 | INDUSTRIAL PRODUCTION INDEX - MATERIALS | 5 |
| IPS34 | INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS | 5 |
| IPS38 | INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS | 5 |
| IPS43 | INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC) | 5 |
| IPS307 | INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES | 5 |
| IPS306 | INDUSTRIAL PRODUCTION INDEX - FUELS | 5 |
| PMP | NAPM PRODUCTION INDEX (PERCENT) | 1 |
| A0m082 | Capacity Utilization (Mfg) | 2 |
| LHEL | INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA) | 2 |
| LHELX | EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF | 2 |
| LHEM | CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA) | 5 |
| LHNAG | CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA) | 5 |
| LHUR | UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA) | 2 |
| LHU680 | UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA) | 2 |
| LHU5 | UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA) | 5 |
| LHU14 | UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA) | 5 |
| LHU15 | UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA) | 5 |
| LHU26 | UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA) | 5 |
| LHU27 | UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS,SA) | 5 |
| A0M005 | Average weekly initial claims, unemploy. insurance (thous.) | 5 |
| CES002 | EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE | 5 |
| CES003 | EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING | 5 |
| CES006 | EMPLOYEES ON NONFARM PAYROLLS - MINING | 5 |
| CES011 | EMPLOYEES ON NONFARM PAYROLLS - CONSTRUCTION | 5 |
| CES015 | EMPLOYEES ON NONFARM PAYROLLS - MANUFACTURING | 5 |
| CES017 | EMPLOYEES ON NONFARM PAYROLLS - DURABLE GOODS | 5 |
| CES033 | EMPLOYEES ON NONFARM PAYROLLS - NONDURABLE GOODS | 5 |
| CES046 | EMPLOYEES ON NONFARM PAYROLLS - SERVICE-PROVIDING | 5 |
| CES048 | EMPLOYEES ON NONFARM PAYROLLS - TRADE, TRANSPORTATION, AND UTILITIES | 5 |
| CES049 | EMPLOYEES ON NONFARM PAYROLLS - WHOLESALE TRADE | 5 |
| CES053 | EMPLOYEES ON NONFARM PAYROLLS - RETAIL TRADE | 5 |
| CES088 | EMPLOYEES ON NONFARM PAYROLLS - FINANCIAL ACTIVITIES | 5 |
| CES140 | EMPLOYEES ON NONFARM PAYROLLS - GOVERNMENT | 5 |
| A0M048 | Employee hours in nonag. establishments (AR, bil. hours) | 5 |
| CES151 | AVG WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WORKERS | 1 |
| CES155 | AVG WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WORKERS | 2 |
| aom001 | Average weekly hours, mfg. (hours) | 1 |
| PMEMP | NAPM EMPLOYMENT INDEX (PERCENT) | 1 |
| HSFR | HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,SA | 4 |
| HSNE | HOUSING STARTS:NORTHEAST (THOUS.U.)S.A. | 4 |
| HSMW | HOUSING STARTS:MIDWEST(THOUS.U.)S.A. | 4 |
| HSSOU | HOUSING STARTS:SOUTH (THOUS.U.)S.A. | 4 |
| HSWST | HOUSING STARTS:WEST (THOUS.U.)S.A. | 4 |
| HSBR | HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR) | 4 |
| HSBNE | HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A | 4 |
| HSBMW | HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A. | 4 |
| HSBSOU | HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A. | 4 |
| HSBWST | HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A. | 4 |
| PMI | PURCHASING MANAGERS' INDEX (SA) | 1 |
| PMNO | NAPM NEW ORDERS INDEX (PERCENT) | 1 |
| PMDEL | NAPM VENDOR DELIVERIES INDEX (PERCENT) | 1 |
| PMNV | NAPM INVENTORIES INDEX (PERCENT) | 1 |

## **Data appendix** (continued...)

| Code | Description | Transf. |
|---|---|---|
| A0M008 | Mfrs' new orders, consumer goods and materials (bil. chain 1982 $) | 5 |
| A0M007 | Mfrs' new orders, durable goods industries (bil. chain 2000 $) | 5 |
| A0M027 | Mfrs' new orders, nondefense capital goods (mil. chain 1982 $) | 5 |
| A1M092 | Mfrs' unfilled orders, durable goods indus. (bil. chain 2000 $) | 5 |
| A0M070 | Manufacturing and trade inventories (bil. chain 2000 $) | 5 |
| A0M077 | Ratio, mfg. and trade inventories to sales (based on chain 2000 $) | 2 |
| FM1 | MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA) | 6 |
| FM2 | MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP(BIL$, | 6 |
| FM3 | MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL$,SA) | 6 |
| FM2DQ | MONEY SUPPLY - M2 IN 1996 DOLLARS (BCI) | 5 |
| FMFBA | MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA) | 6 |
| FMRRA | DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA) | 6 |
| FMRNBA | DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL$,SA) | 6 |
| FCLNQ | COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1996 DOLLARS (BCI) | 6 |
| FCLBMC | WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL$,SAAR) | 1 |
| CCINRV | CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19) | 6 |
| A0M095 | Ratio, consumer installment credit to personal income (pct.) | 2 |
| FSPCOM | S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10) | 5 |
| FSPIN | S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10) | 5 |
| FSDXP | S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM) | 2 |
| FSPXE | S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA) | 5 |
| FYFF | INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA) | 2 |
| CP90 | Commercial Paper Rate (AC) | 2 |
| FYGM3 | INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA) | 2 |
| FYGM6 | INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA) | 2 |
| FYGT1 | INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA) | 2 |
| FYGT5 | INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA) | 2 |
| FYGT10 | INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA) | 2 |
| FYAAAC | BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM) | 2 |
| FYBAAC | BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM) | 2 |
| scp90 | cp90-fyff | 1 |
| sfygm3 | fygm3-fyff | 1 |
| sFYGM6 | fygm6-fyff | 1 |
| sFYGT1 | fygt1-fyff | 1 |
| sFYGT5 | fygt5-fyff | 1 |
| sFYGT10 | fygt10-fyff | 1 |
| sFYAAAC | fyaaac-fyff | 1 |
| sFYBAAC | fybaac-fyff | 1 |
| EXRUS | UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.) | 5 |
| EXRSW | FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$) | 5 |
| EXRJAN | FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$) | 5 |
| EXRUK | FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND) | 5 |
| EXRCAN | FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$) | 5 |
| PWFSA | PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA) | 6 |
| PWFCSA | PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA) | 6 |
| PWIMSA | PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA) | 6 |
| PWCMSA | PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA) | 6 |
| PSM99Q | INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A) | 6 |
| PMCP | NAPM COMMODITY PRICES INDEX (PERCENT) | 1 |
| PUNEW | CPI-U: ALL ITEMS (82-84=100,SA) | 6 |
| PU83 | CPI-U: APPAREL & UPKEEP (82-84=100,SA) | 6 |
| PU84 | CPI-U: TRANSPORTATION (82-84=100,SA) | 6 |
| PU85 | CPI-U: MEDICAL CARE (82-84=100,SA) | 6 |
| PUC | CPI-U: COMMODITIES (82-84=100,SA) | 6 |
| PUCD | CPI-U: DURABLES (82-84=100,SA) | 6 |
| PUS | CPI-U: SERVICES (82-84=100,SA) | 6 |
| PUXF | CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA) | 6 |
| PUXHS | CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA) | 6 |
| PUXM | CPI-U: ALL ITEMS LESS MIDICAL CARE (82-84=100,SA) | 6 |
| GMDC | PCE,IMPL PR DEFL:PCE (1987=100) | 6 |
| GMDCD | PCE,IMPL PR DEFL:PCE; DURABLES (1987=100) | 6 |
| GMDCN | PCE,IMPL PR DEFL:PCE; NONDURABLES (1996=100) | 6 |
| GMDCS | PCE,IMPL PR DEFL:PCE; SERVICES (1987=100) | 6 |
| CES275 | AVG HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS | 6 |
| CES277 | AVG HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS | 6 |
| CES278 | AVG HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS | 6 |
| HHSNTN | U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83) | 2 |