Central Bank & Financial Services Authority of Ireland

5/RT/06 June 2006

Research Technical Paper

(Un)Predictability and Macroeconomic Stability

Antonello D'Agostino
CBFSAI*

Domenico Giannone
ECARES, ULB†

Paolo Surico
Bank of England and University of Bari††

Economic Analysis and Research Department
Central Bank and Financial Services Authority of Ireland
P.O. Box 559, Dame Street
Dublin 2
Ireland
http://www.centralbank.ie

*E-mail: antonello.dagostino@centralbank.ie

†ECARES, Université Libre de Bruxelles - CP 114 - av. Jeanne, 44, B-1050, Brussels, Belgium. E-mail: dgiannon@ulb.ac.be

††Monetary Assessment and Strategy, Bank of England, Threadneedle street - EC2R 8AH - London, United Kingdom. E-mail: paolo.surico@bankofengland.co.uk.

The views expressed in this paper are the personal responsibility of the authors. They are not necessarily held either by the CBFSAI, Bank of England or the ESCB. We are grateful to Jan Groen, Christoph Schleicher, Mark Watson and Karl Whelan for helpful discussions and suggestions. We also thank the seminar participants at CBFSAI internal seminar.

Abstract

This paper documents a new stylized fact of the greater macroeconomic stability of the U.S. economy over the last two decades. Using 131 monthly time series, three popular statistical methods and the forecasts of the Federal Reserve's Greenbook and the Survey of Professional Forecasters, we show that the ability to predict several measures of inflation and real activity declined remarkably, *relative* to naive forecasts, since the mid-1980s. This break down in forecast ability appears to be an inherent feature of the most recent period and thus represents a new challenge for competing explanations of the 'Great Moderation'.

JEL Classification: E37, E47, C22, C53.

Keywords: predictive accuracy, macroeconomic stability, forecasting models, sub-sample analysis, Fed Greenbook.

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 (2005), 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).

The U.S. Great Moderation is one of the most investigated and debated subjects in macroeconomics. The interest derives from the fact that alternative interpretations of the event carry different policy implications (see Bernanke, 2004 for a critical overview of this literature). If the decline in the volatility of inflation and output were simply the result of a more benign macroeconomic environment in the form of smaller non-policy shocks, then nothing could prevent the 1970s from happening again. And, we can only hope that the good luck will persist in the future. On the other hand, if monetary policy was responsible for the large volatility of the 1970s and the beginning of the 1980s, then inspecting the policy decision process could reveal helpful insights to prevent repeating the mistakes of the past.

Despite a large empirical literature, no consensus on the most likely cause of the Great Moderation has yet emerged: the good policy and good luck explanations, for instance, are both consistent with events. This paper presents a new stylized fact which can help discriminate among alternative explanations of U.S. macroeconomic stability. Our main finding is that the historical fall in volatility is associated with a sizable decline in the *relative* predictive accuracy of popular forecasting methods based on large sets of macroeconomic indicators. Using a naive random walk model as benchmark, we show that the fall in forecast ability is, on average, in the order of 30% as measured by Mean Squared Forecast Errors (MSFE). This pattern is not limited to inflation but also extends to several indicators of real economic activity and to short- and long-term interest rates beyond the one month horizon.

The fall in predictive ability 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 ability between naive random walk forecasts and the predictions of those institutions is not rejected for all horizons but the current quarter.

The decline in predictive accuracy 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 sources of the decline in predictability are the dynamic correlations between variables rather than the autocorrelations of output and inflation.

To investigate further the break down in the correlations, we study the ability of each series in the panel of predicting real activity and inflation. On the real side, there is a sizable break in the link between the yield curve and output. This break is concentrated at business cycle frequencies and explains a large fraction of the decline in predictability of forecasting models based on a large number of predictors. On the nominal side, several measures of real activity were the most important predictors of inflation before 1985. During the last two decades, in contrast, no predictor improves upon the naive benchmark model, thereby confirming the out-of-sample break down of the Phillips curve relationship (see Atkenson and Ohanian, 2001).

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 economic agents can effectively anticipate future macroeconomics developments. Econometric models and economists' judgement are however quite helpful for the forecasts of the very short horizon.

Lastly, the literature on forecasting methods, recently surveyed by Stock and Watson (2005a), has devoted a great deal of attention towards identifying the best model for predicting inflation and output. 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 that 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. Section 3 reports the full sample results while the sub-sample analyses are carried out in Sections 4. The latter part shows a robust correlation between greater macroeconomic stability and the decline in the predictability of inflation and real activity that began in the mid-1980s. Section 5 shows that the models of the Fed and other commercial organizations are also associated with a remarkable fall in forecasting accuracy. In Section 6, we find that the break down in the correlation between the slope of the yield curve and output has remarkably contributed to the decline in predictability of real activity. Section 7 concludes and the Appendix reports further sub-sample evidence together with the definitions of the variables.

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 in forecasting 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 Sala, 2005); 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.[1]

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.[2] The Mean Square Forecast Error is

[1]Results are robust to alternative window width selections. For the sake of completeness, we also report in Appendix C the results for the recursive forecasts, which confirm qualitatively the findings in the main text. In the latter case, the estimation period begins always in 1959:1.

[2]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, 2005).

used as metric for evaluating the forecasts, while predictability is defined as the ratio between the MSFEs of a given model and the Naive Random Walk model. A detailed description of the forecasting methods and the data set is reported in Appendix A and Appendix B.

It should be noted that the emphasis of our paper is on the predictability of a given model *relative* to the predictability of a naive model. Another reading of our results is, in fact, that the relative performance of naive forecasts improved during the last two decades. Furthermore, the rise in predictive accuracy of the random walk is simply the flipside of the fall in predictive accuracy of all other models. We use the expression 'predictability of a given model' in *relative* sense throughout the paper, unless otherwise specified.

3 Full-Sample Results

Our analysis begins with the full-sample evidence in Table 1. We report the relative 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.[3]

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

[3]Following Romer and Romer (2000), our inference is based on the regression: $(z_{ht} - \hat{z}_{ht}^{m})^2 - (z_{ht} - \hat{z}_{ht}^{Naive})^2 = c + u_{ht}$ 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. 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.

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.

horizons and for *CPI* and *PCED* at one and three month horizons only. As far as interest rates are concerned, no forecasting model performs significantly better than the naive benchmark.

The evidence in Table 1 is consistent with the results in Stock and Watson (2005a) 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 of alternative candidates. We show in the next section, however, that these findings are driven by the 1970s and the early 1980s when the many macroeconomic series were highly volatile and persistent. This observation appears to limit the benefit of relative performance evaluations over long sample periods that may be subject to parameter instability.

4 Forecast Performance over Sub-Samples

This section presents evidence of a generalized historical decline in the predictability of several measures of inflation and real activity. Results for short- and long-term interest rates are also presented.

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. 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

Table 2 reports the results for all models including the RW. Moving from Period I to Period II, the RW is associated with a sizable moderation in the *absolute* values of the MSFE. The percentage declines of the *relative* MSFEs reported in the last column are sizable, of 40% magnitude on average, and 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 predictive accuracy 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, and as long as *CPI* and *PCE* are concerned 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 (2005b) and D'Agostino and Giannone (2005) about the deterioration of the inflation forecasts on the basis of

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.

Phillips curve models and $FAAR$.

Furthermore, the sizable fall in the persistence and volatility of inflation dated by Cogley and Sargent (2005) and Kim, Nelson and Piger (2004) around the mid-1980s suggests that unpredictability could simply reflect an inherent feature of the most recent observations.

The stylized fact identified in this section does not seem to be limited to the regime shift observed in U.S. monetary policy history. While an international investigation is beyond the scope of this paper, it is interesting to notice that, using a time-varying Bayesian VAR, Benati and Mumtaz (2005) find that inflation in the U.K. has became far less predictable since the introduction of the inflation targeting framework in 1992.

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 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 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. The *FAAR* is the best predictive model in Period I. The significant forecasting advantage of 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. We return to this issue in Section 6.

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 findings of Table 3 are consistent with the results in McConnell and Perez-Quiros (2000), Kim, Nelson and Piger (2004) and Ahmed, Levin and Wilson (2004) of a sizable reduction in the volatility of real activity since the mid-1980s. Furthermore, the evidence presented in this section corroborates the view that the lack of predictability is intrinsic to the post-1985 data rather than specific to a particular forecasting model.

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 relative predictive ability 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.

It is worth to emphasize that ending the earlier sub-sample in 1979:10, which corresponds to the beginning of Volcker's experiment of non-borrowed reserve

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.

targeting, does not overturn the result of interest rate unpredictability at short horizons.[4] Results are available upon request.

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 convincingly

[4] We notice, however, that excluding the period of Volcker's experiment from the earlier subsample improves the forecast ability one year ahead.

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 (2005).

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 nonlinear 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.[5]

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 (FOMC), 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

[5]The data used in this section are available on the web site of the Federal Reserve of Philadelphia. In particular, SPF: http://www.phil.frb.org/econ/spf/spfmed.html; Greenbook: http://www.phil.frb.org/econ/forecast/croushoresdatasets.html; Real-Time: http://www.phil.frb.org/econ/forecast/reaindex.html.

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 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 Predictive Accuracy

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.

The top panel of Table 5 presents the finding for the full postwar period. For inflation, the Greenbook and the SPF forecasts are far more accurate than a

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.

naive model, being associated with significantly lower MSFEs at all horizons. The results of Period I in the middle panel are virtually identical to the full-sample results whereas for the post-1985 period the statistics in the bottom panel paint a quite 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 that more sophisticated forecasts for output are not immune to the generalized decline in predictability.[6]

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

[6] A similar result for SPF predictions on output growth can be found in Campbell (2004). 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.

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			
hor(q)	RW	Fed's Green Book(GB)/RW	Survey of Professional Forecasters(SPF)/RW
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			
hor(q)	RW	Fed's Green Book(GB)/RW	Survey of Professional Forecasters(SPF)/RW
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.

period before 1985. Second, the Greenbook and the SPF forecasts are characterized by a significant decline in the relative predictive accuracy 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.

It is worth to notice, however, that unlike the statistical models during the later subsample the Greenbook retains some advantage over naive forecasts at the short horizon of one quarter. 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.[7]

6 What drives the break down in predictability?

This section further investigates the historical decline in predictability. First, we compute the relative MSFE of 130 bivariate VARs obtained combining each measure of real activity and inflation with the variables in the panel. The results of this exercise are of course only suggestive but can be helpful for shortlisting the suspects. Then, we assess the marginal predictive power of the best predictors

[7]A formalization of these procedures in a data-rich environment can be found in Giannone Reichlin and Sala (2005), and Giannone, Reichlin and Small (2005).

identified by the bivariate VARs analysis. The marginal contribution of selected predictors is measured as the difference between the relative MSFE of the factor model based on the full panel and the relative MSFE of the factor model based on all variables in the panel but the selected predictors.

Table 7 ranks for each period the ten series that best forecast real personal income, industrial production, unemployment rate and employees on nonfarm payrolls one year ahead. Similar findings are obtained using different forecasting horizons, though results are more pronounced for $h = 12$.

Table 7: *Best predictors for real activity*

	Real Personal Income				*Industrial Production*		
	Period I		Period II		Period I	Period II	
RMSFE	predictor	RMSFE	predictor	RMSFE	predictor	RMSFE	predictor
0.66	Spread 3m	0.79	EMP sale trade	0.39	Spread 10y	0.86	CPI med.care
0.67	M2	0.8	Avg w. hours	0.39	Spread 5y	0.91	IP Consumer Gs
0.72	Spread 6m	0.82	Avg w. hours mfg	0.43	Spread 3m	0.92	Dividend Yield
0.75	Spread 5y	0.82	IP Materials	0.45	Spread 6m	0.95	IP Nondur. CGs
0.76	Spread 10y	0.83	IP Durable	0.46	Spread AAA	0.95	Unemp.Dur.:< 5w
0.76	Net Loans	0.84	EMP goods prod.	0.46	Spread 1y	0.95	Vendor deliver
0.79	Spread AAA	0.85	EMP tot private	0.47	M2	0.97	Spread 6m
0.79	CPI med.care	0.86	EMP durable	0.49	Spread BAA	0.97	IP Products
0.8	Spread 1y	0.87	IP Tot index	0.58	Net Loans	0.97	EMP constr.
0.83	Spread BAA	0.88	IP Manufact.	0.70	Dividend Yield	0.97	IP Materials

	Unemployment Rate				*Non-Farm Employees Payrolls*		
	Period I		Period II		Period I	Period II	
RMSFE	predictor	RMSFE	predictor	RMSFE	predictor	RMSFE	predictor
0.47	Spread 10y	0.78	Advertising Index	0.41	Spread 3m	0.50	CPI med.care
0.49	Spread 5y	0.81	Ratio advertising	0.41	Spread 6m	0.52	Spread 3m
0.52	Spread AAA	0.84	EMP constr.	0.42	Spread 10y	0.52	Ratio advert.
0.54	Spread BAA	0.86	EMP goods prod.	0.45	Spread AAA	0.53	Advert. Index
0.54	Spread 1y	0.88	Hours nonagric.	0.46	Spread 5y	0.53	Comm.paper rate
0.54	Spread 6m	0.89	EMP trade et al.	0.47	Spread 1y	0.53	Spread 6m
0.55	Spread 3m	0.89	Spread 1y	0.50	Spread BAA	0.55	Fed funds rate
0.63	M2	0.90	Spread 6m	0.59	Advert. Index	0.55	Spread 1y
0.72	Net Loans	0.90	EMP tot private	0.59	M2	0.56	PCE services
0.75	Stock mkt.	0.91	M2	0.60	Net Loans	0.56	Avg. Hourly earn.

Notes: Spreads are defined as the difference between long term interest rates and the federal funds rate. For the definition of all items, refer to Appendix B. The forecast horizon is one year

The most striking result is that during the pre-1985 period seven measures of interest rate spread are in the top ten list of predictors for real activity.[8] Moreover, virtually all remaining variables in the ranking of Period I are associated

[8] The data set contains 8 interest rate spreads. The difference between the commercial paper rate and the federal funds rate is the only measure of spread with little predictive power for real activity before 1985.

with monetary aggregates, credit conditions or financial market indicators. The shortlist of the later subsample, in contrast, is dominated by measures of real activity and the presence of term spreads becomes the exception.

It is worth to emphasize that the relative MSFEs of the post-1985 sample are systematically larger for real personal income, industrial production, and unemployment rate. Interestingly, moving from the earlier to the later subsample most interest rate spreads lose predictive power, being associated with relative MSFEs far above the statistics of the $10th$ best predictors in Period II. The latter result holds even for employees on nonfarm payrolls, which is the only measure of real activity that can still be accurately forecasted after 1985.

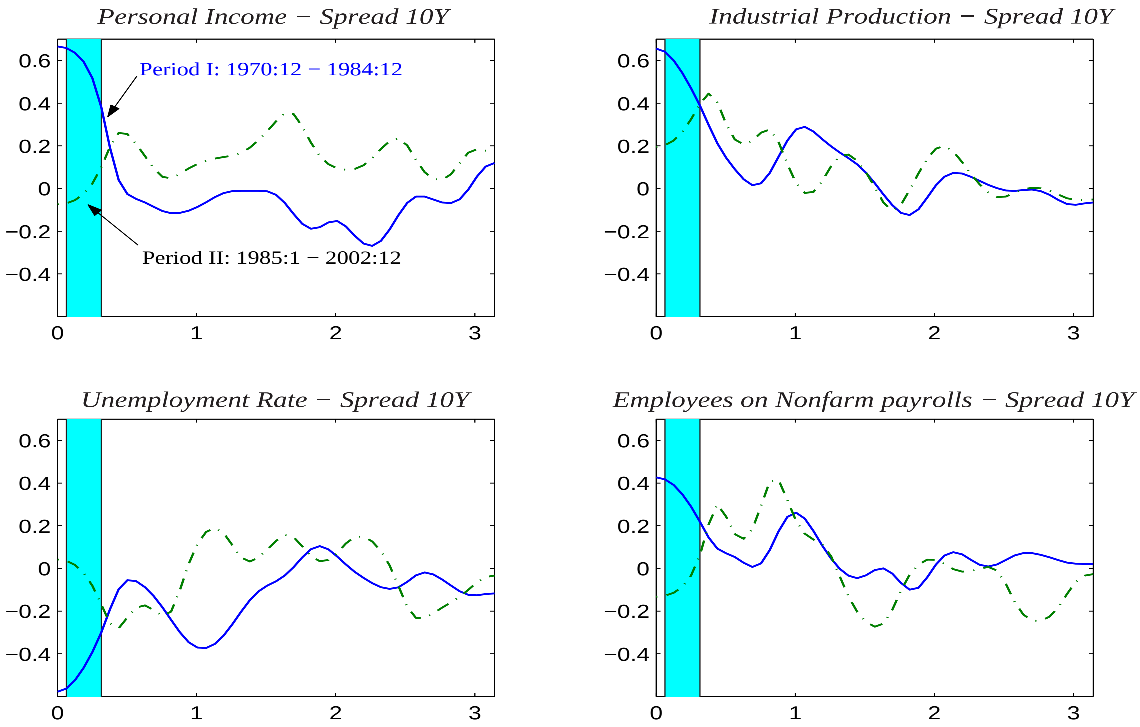
Using bivariate VARs and several measures of asset prices, Stock and Watson (2003c) identify an important break down in the ability of the yield curve of predicting real activity. We show that before 1985 the term spread was, in fact, the best predictor for output among a large number of variables including prices, real activity and money stocks.

Figure 1 further investigates the historical break down in the relationship between the slope of the yield curve and real activity using the dynamic correlation coefficient proposed by Croux, Forni and Reichlin (2001). This measure is attractive for at least two reasons. First, it decomposes the comovement between the term spread and real activity by frequency bands. Second, dynamic correlations take the variance of individual series into account, and thus allow us to compare periods characterized by different volatility.[9]

The dynamic correlations of four measures of real activity and the spread between a 10-year government bond and the federal funds rate are plotted in Figure 1. The results for other term spreads are very similar and available upon request. The vertical axis displays correlation coefficients and the horizontal axis represents frequencies. The solid blue (dotted green) line refers to Period I (Period II) while the gray areas highlight the business cycle frequencies, which correspond

[9]More technically, the dynamic correlation is the real part of coherency and can also be obtained by averaging coherencies at the relevant frequencies. By construction, it takes values in the interval [-1, 1].

Figure 1: *Dynamic correlation by frequency band*



to time periods between 1.5 and 8 years.

The top left panel shows that the most dramatic change between periods occurred at low and business cycle frequencies. The dynamic correlation between the 10-year spread and real personal income moved from values around 0.5 in the pre-1985 sample to coefficients close to -0.2 in the later period. High frequencies, in contrast, are associated with far smaller historical differences and the correlations are, on average, close to zero. A similar picture emerges for the other measures of real activity: large historical differences in dynamic correlations takes place at the business cycle whereas small coefficients characterize higher frequencies. Interestingly, this result also holds for employees on nonfarm payrolls, albeit this measure can still be effectively predicted over the most recent period.

The results of Table 7 and the dynamic correlations in Figure 1 suggest that interest rate spreads may have had strong marginal predictive power for real activity before 1985. To evaluate this hypothesis, Table 8 reports the relative MSFEs of the factor model and the pool of forecasts under two different information sets. For expositional convenience, the left part of Table 8 replicates the results in Table 3, which are based on the full panel of 131 series. The right part of Table 8, in contrast, shows the relative MSFEs of $FAAR$ and $POOL$ based on a panel of 123 series that excludes the term spreads. Numbers in bold denote a fall in predictive accuracy above 10%. The subsample is Period I.

Table 8: *Relative MSFEs in Period I - Real Activity*

FULL PANEL				WITHOUT SPREADS		
Series: Real Personal Income						
hor	FAAR	POOL		hor	FAAR	POOL
1	0.95	0.98		1	0.97	0.99
3	0.86	0.94		3	0.93	0.96
6	0.83	0.96		6	**0.93**	1.00
12	0.84	0.87		12	**0.97**	0.91
Series: Industrial Production						
hor	FAAR	POOL		hor	FAAR	POOL
1	0.65	0.75		1	0.69	0.75
3	0.55	0.73		3	0.59	0.74
6	0.49	0.76		6	**0.59**	0.79
12	0.43	0.72		12	**0.66**	0.77
Series: Unemployment Rate						
hor	FAAR	POOL		hor	FAAR	POOL
1	0.63	0.78		1	0.63	0.78
3	0.52	0.69		3	0.56	0.70
6	0.49	0.75		6	**0.59**	0.78
12	0.56	0.82		12	**0.74**	0.87
Series: Employees on Nonfarm Payrolls						
hor	FAAR	POOL		hor	FAAR	POOL
1	0.51	0.60		1	0.53	0.61
3	0.41	0.53		3	**0.45**	0.54
6	0.42	0.60		6	**0.51**	0.62
12	0.49	0.69		12	**0.67**	0.73

Notes: Period I: sub-sample 1971:1 - 1984:12. The RMSFEs are computed using two information sets. The 'full panel' uses all 131 series. The panel 'without spreads' uses 123 series and excludes the 8 interest rate spreads detailed in Appendix B.

The relative MSFEs of the factor model 'without spreads' are, on average, 10% higher than the MSFEs based on the 'full panel'. In particular, the exclusion of only 8 series produces a sizable fall in the accuracy of the one year ahead forecasts. The decline in predictability is 13% for real personal income, 35% for industrial production, 24% for unemployment rate and 27% for employees on nonfarm payrolls. The deterioration at 6 months horizon is always above 10% whereas at shorter horizons the marginal predictive power of the spreads is of the order of 5%. The relative MSFEs of the pool of forecasts are also systematically larger using the panel 'without spreads', though the differences are less pronounced than for $FAAR$. The fall in forecast ability of $POOL$ is, on average, around 5%, and longer horizons are associated with larger reductions.[10]

In summary, during the 1970s and the first half of the 1980s the slope of the yield curve was a very important predictor of real activity, especially at long horizons.[11] It should be noticed, however, that using a large number of predictors beyond the term spread is still helpful for forecasting real activity before 1985. For inflation, on the other hand, the exclusion of the interest rate spreads does not affect the forecasting performance of $FAAR$ and $POOL$.

The ranking of the bivariate forecasting models for inflation reveals that real variables dominate the shortlists of best predictors over both samples. Moreover, during the last two decades all variables have significantly lost forecasting power and the naive benchmark model turns out to be the best predictor. The break down in the ability of real activity of predicting inflation is consistent with the post-1985 evidence in Atkenson and Ohanian (2001) of a decline in the performances of Phillips curve based forecasting models. The ranking of best predictors for inflation is available upon request.

The finding of a break down in the predictability of real activity and inflation contrasts with the evidence in Ahmed, Levin and Wilson (2004), and Stock and Watson (2003a) of little change in the structure of the economy. If the Great Moderation were only the result of smaller shocks, however, we should have observed little changes in the ability to forecast key macroeconomic variables. Large innovation variances may, in fact, reflect misspecification errors due to the omission of relevant information.

[10] Financial variables and money stocks, in contrast, do not have marginal predictive power for real activity beyond the term spread.

[11] After 1985, the $FAAR$ based on the 'full panel' and the $FAAR$ 'without spreads' produce very similar results.

7 Conclusions

This paper investigates the ability of some widely used econometric models, the Fed's Greenbook and the Survey of Professional Forecasters of predicting 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* experiment could capture. As long as there exists some positive probability that the current macroeconomic stability may come to an end, large policy institutions like Central Banks will have strong incentives to devote resources to forecasting 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.

The generalized decline in the forecast ability of real activity of inflation is associated with the decline in the volatility of inflation and real activity documented by McConnell and Perez-Quiros (2000), and Cogley and Sargent (2005). While investigating any possible causality is beyond the scope of the paper, it is intriguing to notice the timing of events as the break dates of these two stylized facts are concentrated in the first half of the 1980s. And, the one month ahead predictions of the short- and long-term interest rates are the only more accurate forecasts over the most recent period. This observation suggests that a more transparent communication strategy and better monetary policy management could have contributed to interest rate forecast ability.

At a more general level, this paper presents a new fact of U.S. greater macroeconomic stability. The important implication that can be drawn from our analysis is that any theoretical model aimed at explaining the 'Great Moderation' must be capable of accounting for the historical decline in the ability of predicting inflation and real activity. An interesting avenue for future research is to investigate the correlation between forecast ability and policy regimes. Ironically, it may well be the case that the fall in predictability of inflation and output reflects improvements in monetary policy.

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(3), 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 Haroon Mumtaz (2005), "The 'Great Stability' in the U.K.: Good Policy or Good Luck?", mimeo, Bank of England.

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

[7] Boivin, Jeanne and Marc Giannoni (2005), "Has Monetary Policy Become More Effective?", *The Review of Economic and Statistics*, forthcoming.

[8] Campbell, Sean, D., (2004), "Macroeconomic Volatility, Predictability and Uncertainty in the Great Moderation: Evidence from the Survey of Professional Forecasters", Federal Reserve Board Finance and Economics Discussion Series No. 2005-52.

[9] Cogley, Timothy W. and Thomas Sargent, (2005) "Drifts and Volatilities: Monetary Policies and Outcomes in the Post World War II U.S.," *Review of Economic Dynamics* 8, pp. 262-302.

[10] Croux, Christophe, Mario Forni and Lucrezia Reichlin, "A Measure of Comovement for Economic Variables: Theory and Empirics," *The Review of Economic and Statistics* 83, pp. 232-241.

[11] D'Agostino, Antonello and Domenico Giannone, (2005), "Comparing Alternative Predictors Based on Large-Panel Factor Models", mimeo, ECARES, Universite Libre de Bruxelles.

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

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

[14] Giacomini, Raffaella, and Halbert White, (2005), "Tests of Conditional Predictive Ability", mimeo, UCLA.

[15] Giannone, Domenico, Lucrezia Reichlein and Luca Sala, (2005), "Monetary Policy in Real Time ", *NBER Macroeconomics Annual*, Eds: Mark Gertler and Ken Rogoff, forthcoming.

[16] Giannone, Domenico, Lucrezia Reichlin and David Small, (2005), "Nowcasting GDP and Inflation: The Real Time Informational Content of Macroeconomic Data Releases", Federal Reserve Board, Finance and Economics Discussion Series, 2005-42.

[17] 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, 608 – 616.

[18] Kim, Chang-Jin, Charles R. Nelson and Jeremy Piger (2004), "The Less-Volatile U.S. Economy: A Bayesian Investigation of Timing, Breadth, and Potential Explanations", *Journal of Economic and Business Statistics*, 22, 80-93.

[19] 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(5), 1464 – 1476.

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

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

[22] Rudebusch, Glenn, (2002) "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia", *Journal of Monetary Economics* 49(6), 1161-1187.

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

[24] Stock, James and Mark Watson (2005b), Has Inflation Become Harded to Forecasts?, mimeo Harvard University.

[25] 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.

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

[27] Stock, James and Mark Watson (2003c) "Forecasting Output and Inflation: the Role of Asset Prices" Journal of Economic Literature vol. 41, pp. 788-829.

[28] West, Kenneth, D. (1996) "Asymptotic Inference about Predictive Ability" *Econometrica* 64, 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

$$Y^h_{i,t+h} = \alpha^{h,Naive}_i + e^{h,Naive}_{i,t+h}$$

Autoregressive

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

Augmented distributed lag

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

r-factor model

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

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 for price series.

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	$Z_{t+h}-Z_t$	ΔZ_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	$Z_{t+h}-Z_t$	ΔZ_t
10-Yr Tbond Rate	TBOND	$Z_{t+h}-Z_t$	ΔZ_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 $W_t = (W_{1t}, ..., W_{nt})', t = T_{0T}, ..., T$, 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:

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

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

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

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

Pooled forecasts from different ADL models are computed as:

$$\hat{Y}^h_{i,T+h|T}(POOL) = \frac{1}{(n-1)} \sum_{j \neq i} \hat{Y}^h_{i,T+h|T}(ADL_j)$$

For rolling sample estimates we consider observations from a fixed window of ten years,. i.e. as data are monthly, $T_{0T} = T - 120$. For recursive samples, we always have T_{0T} =January 1959.

Our Mean Square Forecast error measure for forecast evaluation is equal to:

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

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. Predictability is defined as the ratio between the MSFE of a given model and the naive model:

$$PRED^{t_1}_{t_0}(i,h,m) = \frac{MSFE^{t_1}_{t_0}(i,h,m)}{MSFE^{t_1}_{t_0}(i,h,Naive)}$$

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

Appendix C: Recursive Sub-samples

Table C1: Relative MSFEs - Full Period using Recursive Samples

Random Walk (absolute values)

hor	PPI	CPI	PCED	PI	IP	UR	EMP	TBILL	TBOND
1	0.45	0.11	0.06	45.42	75.00	0.03	9.28	0.31	0.11
3	1.79	0.57	0.31	13.80	44.88	0.14	6.96	1.27	0.47
6	4.24	1.54	0.89	7.56	33.35	0.42	6.23	2.43	0.97
12	11.15	4.58	2.71	4.88	23.63	1.24	5.20	4.46	2.10

Method AR (relative to RW)

hor	PPI	CPI	PCED	PI	IP	UR	EMP	TBILL	TBOND
1	1.00	0.89	0.86	1.03	0.83	0.91	0.57	0.96	0.90
3	1.02	0.92	0.86	1.03	0.82	0.80	0.50	1.10	1.06
6	1.00	0.88	0.85	1.00	0.89	0.83	0.59	1.06	1.02
12	1.06	0.96	1.01	0.98	0.94	0.93	0.75	1.16	1.00

Method FAAR (relative to RW)

hor	PPI	CPI	PCED	PI	IP	UR	EMP	TBILL	TBOND
1	0.94	0.80	0.80	0.96	0.71	0.73	0.48	0.89	0.89
3	0.85	0.71	0.78	0.92	0.60	0.57	0.39	0.98	1.07
6	0.76	0.62	0.77	0.89	0.59	0.52	0.43	0.86	1.03
12	0.71	0.62	0.83	0.90	0.59	0.55	0.53	0.90	1.03

Method POOL (relative to RW)

hor	PPI	CPI	PCED	PI	IP	UR	EMP	TBILL	TBOND
1	0.97	0.86	0.85	1.01	0.78	0.85	0.53	0.93	0.89
3	0.95	0.85	0.82	0.99	0.74	0.72	0.45	1.07	1.06
6	0.90	0.79	0.80	0.94	0.77	0.72	0.52	0.99	1.02
12	0.92	0.80	0.90	0.92	0.77	0.78	0.66	1.08	1.01

Table C2: Relative MSFEs - Sub-Periods using Recursive Samples

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

Series: Producer Price Index

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	0.55	1.01	0.94	0.97	1	0.37	0.98	0.95	0.97
3	2.20	1.04	0.78	0.93	3	1.47	1	0.92	0.97
6	5.63	0.97	0.63	0.84	6	3.15	1.04	0.94	0.99
12	16.98	1.06	0.60	0.88	12	6.59	1.08	0.94	1

Series: Consumer Price Index

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	0.17	0.83	0.74	0.79	1	0.07	0.99	0.91	0.98
3	0.91	0.86	0.62	0.78	3	0.30	1.06	0.91	1.02
6	2.68	0.81	0.51	0.71	6	0.65	1.12	0.97	1.04
12	8.58	0.89	0.49	0.72	12	1.45	1.29	1.22	1.2

Series: Personal Consumption Expenditure Deflator

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	0.08	0.73	0.69	0.71	1	0.05	1.05	0.96	1.03
3	0.49	0.74	0.67	0.70	3	0.17	1.13	1.01	1.09
6	1.56	0.74	0.67	0.69	6	0.37	1.22	1.1	1.17
12	5.16	0.94	0.73	0.82	12	0.78	1.36	1.35	1.28

Series: Real Personal Income

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	38.35	1.04	0.95	1	1	50.94	1.03	0.97	1.01
3	16.89	1.04	0.89	0.98	3	11.39	1.02	0.95	1.00
6	10.15	1.02	0.8	0.94	6	5.53	0.97	1.02	0.94
12	6.80	0.99	0.8	0.89	12	3.38	0.97	1.06	0.96

Series: Industrial Production

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	123.40	0.80	0.64	0.74	1	37.12	0.93	0.91	0.89
3	79.90	0.83	0.53	0.74	3	17.48	0.79	0.82	0.75
6	59.27	0.91	0.46	0.77	6	13.07	0.82	1.03	0.79
12	41.45	0.95	0.41	0.75	12	9.69	0.89	1.23	0.85

Series: Unemployment Rate

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	0.05	0.87	0.64	0.80	1	0.02	0.98	0.91	0.93
3	0.24	0.78	0.52	0.69	3	0.06	0.86	0.70	0.80
6	0.74	0.83	0.48	0.72	6	0.16	0.80	0.69	0.74
12	2.19	0.95	0.49	0.79	12	0.49	0.85	0.75	0.78

Series: Employees on Nonfarm Payrolls

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	15.88	0.62	0.49	0.58	1	4.11	0.40	0.46	0.38
3	11.72	0.57	0.40	0.51	3	3.24	0.29	0.35	0.27
6	10.30	0.67	0.41	0.59	6	3.05	0.37	0.48	0.35
12	8.24	0.82	0.46	0.70	12	2.81	0.60	0.7	0.58

Series: 3 Months Treasury Bills

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
5 1	0.63	0.97	0.9	0.94	1	0.0	0.86	0.71	0.81
3	2.54	1.10	0.98	1.07	3	0.28	1.11	0.98	1.03
6	4.48	1.02	0.81	0.96	6	0.83	1.20	1.08	1.12
12	7.01	1.21	0.85	1.12	12	2.46	1.04	1.01	0.99

Series: 10 Years Treasury Bonds

hor	RW	AR	FAAR	POOL	hor	RW	AR	FAAR	POOL
1	0.16	0.91	0.91	0.91	1	0.07	0.87	0.86	0.86
3	0.67	1.10	1.11	1.10	3	0.32	1	1.02	1.00
6	1.23	1.02	1.03	1.02	6	0.76	1.02	1.04	1.02
12	2.37	0.99	1.03	1	12	1.89	1.01	1.04	1.01