Diverse Beliefs, Forecast Errors and Central Bank Policy

by

July 12, 2005

Abstract. The Fed and private forecasters have not forecasted inflation and GDP growth accurately and have assessed incorrectly the effects of monetary policy on these variables. The mistakes are quantitatively different across three monetary regimes studied but, surprisingly, we find significant *qualitative* similarity in the mistaken assessment of the stabilization effects of monetary policy. Over the period of study (1965:11-1995:11) we find that

(i) for long horizons the Fed and private inflation forecasts (i.e. over 3 quarters) either underestimated or overestimated future developments: when forecasting low inflation, forecasts were too low and when forecasting high inflation, they were too high. The same pattern is true for GDP growth but *over all horizons;*

(ii) the Fed inflation forecasts *over-estimated* the stabilization effects of monetary policy over short horizons and *under-estimated* these effects over long horizons;

(iii) all forecasters of GDP growth underestimated, for almost all horizons, the stabilization effect of interest rates on GDP growth in all three monetary regimes. Forecasters underestimated both the stimulating effects of low interest rates as well as the inhibiting effect of high interest rates on GDP growth. We conjecture this persistent error is a consequence of the academic bias towards money neutrality during the period of study.

The quality of the Fed's forecasts of inflation and GDP growth is sufficiently poor to conclude that monetary policy decisions which are unexpected by the private sector should be confined to special circumstances and to emergencies such as wars or financial crises. We present extensive evidence in support of the perspective that private forecasters and the Fed use different models to forecast inflation and GDP growth: heterogeneity of forecasting models is the norm in the marketplace. We study the similarity and differences of such models.

JEL classification: E17; E27; E3, E42; E47; E5; E52;E58.

Keywords: Federal Reserve; unexpected decisions; monetary policy; forecast error; forecasting models; inflation; GDP growth ; overestimation; underestimation; heterogenous models; Rational Beliefs.

Diverse Beliefs, Forecast Errors and Central Bank Policy[1]

by

July 12, 2004

1. Introduction

The literature on policy discretion assumes the central bank selects an optimal mix of inflation and unemployment (e.g. Kydland and Prescott (1977), Barro and Gordon (1983)). Such formulation considers policy discretion an *incentive problem* based on three assumptions. First, the central bank has private information advantage over the private sector. Second, the bank knows private expectations which it then takes as given. Third, the bank and the private sector know the impact of central bank policy actions. Does the Fed have private information? Romer and Romer (2000) and Ellingsen and Söderström (2001) accept the very limited evidence as valid. Alternative views are expressed by Cochrane and Piazzesi (2002), Faust, Swanson and Wright (2004), Kurz (2002) and Sims (2002). The Fed has provided all data used in preparing its Green Book forecasts, showing the information comes only from public sources available to all. Sims (2002) shows the Fed forecasts are superior to private forecasts only at some intervals. Most economists doubt the Fed's staff uses private information to forecast macroeconomic variables since it is hard to see what that information could be. Faust, Swanson and Wright (2004) suggest private information might be the bank's own policy surprises but then they show the "release" of such information does not improve private forecasts. Given these considerations we note that the Fed's Green Book forecasts, studied in this paper, *are prepared by the staff* for the Open Market Committee *before* actual policy decisions are made. Hence, Fed forecasts in this paper *are not* forecasts of the FOMC.

Does the Fed and private sector know the impact of a central bank policy? The literature on policy discretion assumes the effect of a bank's action is instantaneous and known. In reality it takes

[1] This research was supported by a grant of the Smith Richardson Foundation to the Stanford Institute for Economic Policy Research (SIEPR). The author is indebted to Christina D. Romer and to David H. Romer for their generosity in providing us with their initial data file, to Randell Moore of the Blue Chip Economic Indicators for providing the detailed BLU data used in this paper and to Dean Croushore for help with respect to the SPF and Federal Reserve data files. The author thanks Min Fan, Sherrill Shaffer, Eric Swanson and Ho-Mou Wu for helpful discussions and many comments on earlier drafts. He thanks Andrea Beltratti, Jean Boivin, Arturo Ricardo Gonzalez de la Mota, Maurizio Motolese and Carsten Nielsen for helpful comments. He also thanks Albert Chun, Scott Fulford, Hehui Jin and Peyron Law for assistance with the technical aspects of the work.

time for a bank's action to impact inflation and output growth and the effect is uncertain. Indeed, central banks do not choose output growth and inflation rates but only nominal interest rates, hoping to impact future inflation and output. To implement any policy the Fed must forecast accurately the consequences of alternative policy actions. But, when agents do not know the true structure of the economy and have diverse beliefs about it, they also have diverse beliefs about the effects of a Fed's action. Kurz (1997) and Kurz, Jin and Motolese (2005a) (2005b) show that diversity and dynamics of beliefs are key determinants of market volatility. Hence, the effect of a Fed's action depends upon private expectations and these are not known to the bank with precision. In short, in a world with diverse beliefs and where the Fed and the private sector do not know the exact effects of a Fed's policy action, it is far from clear how the Fed can take advantage of private expectations. Moreover, what may appear as the Fed's private information is nothing but the Fed's subjective forecasting model based on its own interpretation of public information.

We thus have two views of discretion. Under the three assumptions above it is an incentive problem arising when the underlying structure is known. But when these assumptions do not hold, the concept of "discretion" is ambiguous. For the private sector discretion entails unexpected decisions, constituting an unpredictable component of the policy. For this reason we use in this paper the neutral terminology of "unexpected decisions" where "unexpected" refers to the perspective of the private sector. In either case the necessary condition to justify Fed's unexpected decisions is the bank's ability to forecast accurately the effect of its actions. The main question asked in this paper is then simple: is the Fed able to forecast inflation and GDP growth accurately enough to justify the use of policy surprises and can the Fed accurately assess the effect of its policy actions on future inflation and GDP growth? In addition, how do private forecasters perform compared with the Fed's forecasts? The issue at hand is then the quality of the Fed's and private forecasting.

We evaluate forecast accuracy by studying predictability of forecast error functions. This may be viewed as related to the literature on forecast rationality and tests of the Rational Expectations hypothesis. *This is not our perspective* hence we do not cite that literature. In many papers (e.g. Kurz (1994), (1997), Kurz, Jin and Motolese (2005a) (2005b)) we have argued that Rational Expectations are incompatible with the observed level of real and financial market volatility hence it is merely a mathematical reference point. Our hypothesis is that the economy's dynamics is non stationary with time varying technology, institutions and monetary policy regimes. In such an economy agents cannot learn from data the economy's true stochastic structure. Without this knowledge agents must use subjective probability models and these are generally not equal to the true probability under which the data is generated. In that case forecast errors are orthogonal to histories of observables with probability one only under subjective probabilities, not under the true probability. Hence, as a general proposition we should expect that orthogonality with respect to the observables *fail most of the time*. But then a failure of orthogonality does not imply irrational behavior in any sense since agents do the best with what they know. Indeed, we examine the forecasting mistakes made within each specific era and such a retrospective study does not imply that, at the time, forecasters could have done any better. We further explore this issue in Section 4.2.

We briefly sum our results. The Fed and private sector have a poor record of forecasting inflation and GDP growth and they have assessed incorrectly the effects of monetary policy on these variables. The pattern of errors was quantitatively different across the three monetary regimes we studied. We later explain why we should expect the pattern of such errors to vary over time. However, we find pattern of *qualitative* similarity of forecasting errors over 1965-1995:

(i) at long horizons (i.e. over 3 quarters) inflation forecasts underestimated or overestimated events: when forecasting low inflation, forecasts were too low and when forecasting high inflation they were too high. We later offer a conjecture on how to explain the persistence of this pattern of errors;

(ii) the Fed's inflation forecasts *over-estimated* the effect of tight monetary policy over short horizons and *under-estimated* the effect of such policy over long horizons: the leg in response to changes in the Funds rate was longer than expected!

(iii) all forecasters of GDP growth underestimated, for almost all horizons, the stimulating effects of low interest rates and the dampening effect of high rates on GDP growth. A bias is exhibited by all forecasters in underestimating the real stabilization effect of monetary policy on GDP growth. This result is particularly surprising since it is persistent over the entire period of study.

The quality of the Fed's forecasts of inflation and GDP growth is sufficiently poor that we conclude monetary policy surprises should be confined to special emergency circumstances since Fed's policy surprises add to private uncertainty. We present evidence that private forecasters and the Fed *use different models* and heterogeneity of forecasting models is the market norm. Neither the Fed nor private forecasters know the true data generating mechanism of the economy.

2. The Data

2.1 *The Mean and Median Forecast Data*

We use the Fed's Green Book forecasts from 1965:11 through 1995:11 and three private data files. First, the Blue Chip Economic Indicators (BLU) who provide forecasts of about 50 large corporations, financial institutions and consulting firms. These forecasts were made from 1980:1 through 1995:11, covering many economic variables. The second is from the Survey of Professional Forecasters (SPF), reporting forecasts of private forecasters, currently conducted by the Federal Reserve Bank of Philadelphia. SPF continues the American Statistical Association\NBER survey started in 1968. Hence, the BLU and SPF surveys contain forecasts of a large number of different forecasters over many years. One difference between BLU and SPF is that the identity of the SPF forecasters is not revealed while the identity of the BLU forecasters is public. The third source are forecasts made by Data Resources, Inc. (DRI) starting in 1970:7. Hence, DRI data consist of time series of *one forecaster* while BLU and SPF files consist of heterogeneous collections of forecasts. Romer and Romer (2000) used this data up to 1991:11 for the Green Book and DRI forecasts. They also used the mean (i.e. consensus) forecasts of BLU and median forecasts of SPF as forecast data up to 1991:11. We received this file from C. and D. Romer and updated all data files to 1995:11. A detailed description of these data is available in Romer and Romer (2000) pages 430-433. Romer and Romer (2000) also discuss in detail the problem of data comparability due to different dates of release. We make several comments to help the reader understand the nature of the data used.

(i) Frequency of Data. Data sources are released at different frequencies. The Fed's Green Book forecasts are prepared by the professional staff for the Open Market Committee since 1965:11 and these forecasts are kept confidential for five years. Hence, the Fed's forecasts are not available to private forecasters when they make their forecasts. In the 1960's and 1970's the committee met almost each month. Since the 1980's it has typically met eight times a year. Since there are no forecasts in months when the committee does not meet, the frequency of the Fed's forecasts has changed with the frequency of the committee's meetings. The BLU forecasts are made around the third day of each month and circulated to subscribers around the tenth day of each month starting with 1980:1. The forecasts of Data Resources, Inc. are issued three times at each quarter: one early, one in the middle and one late in the quarter since 1970:7 but the middle forecasts are available only since 1980. The SPF survey of forecasters is conducted near the end of the second month of each quarter and all SPF data is quarterly. To attain maximal matching of data sources our forecasting files are arranged so that, for every forecast horizon, the data is constructed as monthly data with missing observations. The BLU data are frequent but cover only the period 1980:1 - 1995:11. The SPF forecasts for inflation are available since 1968:11 and for real GDP growth since 1981:8 but the main limitation imposed by SPF is the fact that it is available only four times each year.

(ii) Forecast Horizon. The actual inflation rate and growth rate of GDP in the quarter in which the forecasts are made are not known. Hence, each set of forecasts includes a forecast for the "current quarter" and we denote this horizon by h = 0. Hence, h = 1 means "the quarter following the quarter in which the forecasts were made." The Green Book horizon goes up to seven future quarters but it varies over time. The BLU consensus forecasts are available for six and sometimes for seven future quarters. The DRI forecast horizon is typically seven future quarters and the SPF is typically four future quarters. The number of observations available for long horizon forecasts is typically small and equations estimated for long horizons using multiple sources are very unreliable. Sometimes we just avoid the seven quarter horizon models due to insufficient data.

(iii) Actual data on inflation and GDP growth. Data revision has been discussed in the literature as a complicating factor since it raises the question of what forecasters are forecasting. The initial GDP statistics, released about 45 days after the end of each quarter, are incomplete. These initial estimates contain significant errors since some component series are not available, hence revisions are needed. The first revision is released at the end of the second month following each quarter and the second revision at the end of the third month following each quarter (that is, at the end of the subsequent quarter). Further revisions are made each July and reevaluation each five years. We study in this paper all forecasting horizons hence we need a reasonably complete measure of inflation and growth which is conceptually close to what the agents were forecasting. We thus use as our realized data the second revision of inflation rates measured by the GDP deflator (GDP after 1991:11) and of real GDP growth rates. In our view the second revision is complete and is free of any conceptual reworking. It is thus conceptually close to what the agents are forecasting[2].

(iv) Serial Correlation. Serial correlation in forecast errors is inevitable (for details, see Romer and Romer (2000) page 433) and suggests that a correction is in order. We use the Newey and West (1987) procedure to compute robust standard errors in all equations of this paper.

2.2 *Distributions of BLU and SPF Forecast Data*

We also use panel data of *individual* forecasts of inflation and real GDP growth reported by SPF and BLU. The SPF Survey assigns a number to each participating forecaster. Participants drop out over time and new forecasters are added, but a record of changes in the identity of participants is kept. Frequency and horizons of the distributional SPF data is the same as the SPF median forecast data but the number of forecasters varies with time. The situation with BLU is different.

BLU reports each month the consensus forecasts for the current quarter and for six or seven quarter into the future. It also discloses the names of forecasters who provide forecasts of various economic variables for the "Current Year" and for "Next Year." In other words, we have monthly data on forecasts for two full calendar years: one in which the forecasts are made and the second is the calendar year which follows. Hence, in any year we get 12 pairs of different forecasts of various economic variables for the same two full calendar years. These individual forecasts cover the period from 1980:1 through 2001:5.

2.3 *Illustrations of Forecast Heterogeneity*

It is instructive to illustrate first the heterogeneity of forecasts in the market. Table 1A presents the January 1991 BLU individual forecasts of GDP growth and inflation for 1991. About half of the forecasters predicted that 1991 will be a recession year while the other half disagreed. The actual growth rate in 1991 was -.5% and the inflation rate was 3.6%.

Now, place yourself in January 1991 and make a stationary forecast of GDP growth without judgment about conditions in 1991. This requires a time invariant model, conditioned on past data to

[2] The Romer and Romer (2000) file stops in 1991:11 partly because they did not have the Green Book data and partly because of the switch from GNP to GDP in government statistics. In extending the data to 1995:11 we accepted possible distortion that might be created by this data switch. We think it is minor compared with forecast errors analyzed in this paper.

forecast growth rates of GDP. We examined the question with many variants of a Stock and Watson (1999a),(1999b),(2001) model and estimated it, in accord with their specifications, with diffusion indexes and averaged bivariate VAR forecasts utilizing a large number of U.S. time series. All non-judgmental forecasts of GNP growth in January 1991 were *higher than most private forecasts.*

Table 1A: January 1991 BLU Forecasts of GDP Growth and Inflation for 1991

Forecasted Percent Change In January 1991 Forecast for 1991	Real GDP	GDP Price Deflator
Sears Roebuck & Co.	1.6H	4.2
Amhold & S. Bleichroeder	1.2	4.8
Prudential Bache	1.2	3.3L
Chicago Corporation	1.1	4.1
Bostian Economic Research	1.0	4.0
Faimodel	1.0	3.7
Cahners Economics	0.9	4.3
Wayne Hummer & Co. - Chicago	0.8	4.3
Nat'l. City Bank of Cleveland	0.7	4.6
Inforum - Univ. of Maryland	0.7	3.8
CRT Government Securities	0.6	4.0
Dun & Bradstreet	0.6	4.0
Conference Board	0.5	4.7
Econoclast	0.5	4.0
First National Bank of Chicago	0.5	3.8
Univ. of Michigan M.Q.E.M.	0.4	4.7
Manufacturers Natl. Bank - Detroit	0.3	4.5
Turning Points (Micrometrics)	0.2	4.3
Brown Brothers Harriman	0.2	4.0
Dean Witter Reynolds, Inc.	0.1	4.0
LaSalle National Bank	0.1	3.6
Northern Trust Company	0.0	4.3
Evans Economics	0.0	4.0
Morris Cohen & Associates	-0.1	5.0H
Prudential Insurance Co.	-0.1	4.5
Chrysler Corporation	-0.1	4.1
Econoviews International Inc.	-0.1	3.9
U.S. Trust Co.	-0.2	4.3
Reeder Associates (Charles)	-0.3	4.9
Siff, Oakley, Marks, Inc.	-0.3	4.8
Morgan Stanley & Co.	-0.3	4.7
Eggert Economic Enterprises, Inc.	-0.3	3.9
CoreStates Financial Corp.	-0.4	4.3
Mortgage Bankers Assn. of America	-0.4	4.3
Bank of America	-0.4	3.6
E.I. Du Pont de Nemours & Co.	-0.5	4.8
National Assn. of Home Builders	-0.5	4.5
Metropolitan Life Insurance Co.	-0.5	4.5
Ford Motor Company	-0.6	4.6
Chase Manhattan Bank	-0.6	4.0
U.S. Chamber of Commerce	-0.7	5.0H
Manufacturers Hanover Trust Co.	-0.7	4.4
Bankers Trust Co.	-0.7	4.4
Laurence H. Meyer & Assoc.	-0.7	4.0
Security Pacific Nat'l. Bank	-0.7	4.0
PNC Financial Corp.	-0.9	4.3
UCLA Business Forecast	-0.9	4.2
Merrill Lynch	-1.1	4.4
Georgia State University	-1.1	3.6
Equitable Life Assurance	-1.2	4.7
Morgan Guaranty Trust Co.	-1.2	3.8
Shawmut National Corp.	-1.3L	4.0

Table 1B exhibits information for May 2000. GDP growth rate in 2000 was 4.1% and the inflation rate 2.3%. It is surprising that in May of 2000, five months into the year, large variability in forecasts is present and almost all GDP forecasts turned out wrong. We again made non-judgmental forecasts and this time the stationary forecasts of GDP growth turned out *lower than most private forecasts.* We thus see that the distribution of forecasts fluctuates over time around the stationary forecasts. In 1991 private forecasters were conservative, forecasting below the stationary forecast while in May of 2000 their judgment was aggressive, forecasting above the stationary forecast.

Table 1B: May 2000 Blue Chip Forecasts of GDP Growth and Inflation for 2000

Forecasted Percent Change In May 2000 Forecast for 2000	Real GDP	GDP Price Deflator
First Union Corp.	5.3H	2.0
Turning Points (Micrometrics)	5.2	2.1
J P Morgan	5.2	2.1
Evans, Carroll & Assoc.	5.1	2.2
Mortgage Bankers Assn. of Amer.	5.1	2.1
Goldman Sachs & Co.	5.1	2.1
U.S. Trust Co.	5.1	2.0
US Chamber of Commerce	5.1	2.0
Banc of America Corp.	5.1	2.0
Morgan Stanley Dean Witter	5.1	1.9
Wayne Hummer Investments LLC	5.0	2.3
Bank One	5.0	2.1
Nomura Securities	5.0	1.9
Merrill Lynch	5.0	1.9
Perna Associates	4.9	2.3
National Assn. of Home Builders	4.9	2.1
Macroeconomic Advisers, LLC	4.9	2.1
Prudential Securities, Inc.	4.9	2.0
LaSalle National Bank	4.8	2.3
Conference Board	4.8	2.3
Wells Capital Management	4.8	2.2
DuPont	4.8	2.1
Northern Trust Company	4.8	2.1
Chicago Capital, Inc.	4.8	2.0
Deutsche Bank Securities	4.8	1.8
Chase Securities, Inc.	4.8	1.8
Credit Suisse First Boston	4.8	1.8
Comerica	4.7	2.4
Moody's Investors Service	4.7	2.2
Fannie Mae	4.7	2.0
Federal Express Corp.	4.7	2.0
SOM Economics, Inc.	4.7	1.9
National Assn. of Realtors	4.7	1.9
National City Corporation	4.7	1.9
Clear View Economics	4.7	1.9
Eggert Economic Enterprises, Inc.	4.6	2.1
WEFA Group	4.6	1.9
Eaton Corporation	4.6	1.9
Bear Stearns & Co., Inc.	4.6	1.2 L
Ford Motor Company	4.5	1.8
Motorola	4.5	1.7
Standard & Poors Corp.	4.5	1.7
UCLA Business Forecasting Proj.	4.4	2.1
Inforum - Univ. of Maryland	4.4	2.0
Prudential Insurance Co.	4.4	1.9
Weyerhaeuser Company	4.3	2.2
DaimlerChrysler AG	4.3	2.0
Georgia State University	4.2	2.2
Kellner Economic Advisers	4.2	2.0
Econoclast	4.1	2.0
Naroff Economic Advisors	4.0 L	2.5 H

Examination of forecast distributions in 1980 - 2003 reveals that Tables 1A-1B are typical. Major financial institutions have access to the same information when making forecasts of inflation and GDP growth. Yet, they make diverse forecasts at any date and their rankings within the forecast distribution *fluctuate dramatically over time*. The stationary econometric forecasts is often different from the median forecast since the *distribution itself fluctuates* over time. One concludes that these forecasters use different models to interpret the same public information used in making forecasts.

4. Testing for Accuracy of Inflation and GDP Growth Forecasting

Let us start by specifying some notation which we use:

$\pi_{t,h}$ - annualized inflation rate *h quarters after date t*. Dates are measured in quarters.

$g_{t,h}$ - real annualized growth rate of GDP *h quarters after date t*. h = 0 means "current" quarter.

$\pi^k_{t,h} \equiv E^k[\pi_{t,h}|I_t]$ forecast of $\pi_{t,h}$ by k at date t, given information I_t, k = GB, BLU, SPF and DRI.

$g^k_{t,h} \equiv E^k[g_{t,h}|I_t]$ forecast of $g_{t,h}$ by k at date t, given information I_t, k = GB, BLU, SPF and DRI.

We investigate the structure of forecast errors of the Fed *and* private forecasters. However, we start with an interesting preliminary step which examines the improvements in accuracy of the Fed forecasts by utilizing data on private forecasts, known to the of the Fed at the time of forecasts.

4.1 *Preliminary Note: Private Forecasts Improve the Performance of the Fed's Forecasts*

When staff members of the Fed prepare their forecasts they know the forecasts of private forecasters. If the Fed staff knew the true probability law of the stochastic process of inflation and GDP growth and if they used all available information, it follows from properties of conditional probability that private forecasts are orthogonal to the forecast errors of the Fed and should contribute nothing to the Fed forecasts. To study if the use of private forecasts can improve the Fed's forecasts we first examine the simple framework used by Romer and Romer (2000). It formulates the relation between forecasts and realizations of inflation and GDP growth in the form

$$\pi_{t,h} = \alpha_0^{h,GB} + \alpha_1^{h,GB}\pi_{t,h}^{GB} + \varepsilon_{t,h}^{GB} \tag{1a}$$

$$g_{t,h} = \delta_0^{h,GB} + \delta_1^{h,GB} g_{t,h}^{GB} + \vartheta_{t,h}^{GB} . \tag{1b}$$

We first estimate the parameters of (1a)-(1b) using data for 1965:11-1995:11. These estimates are the basic reference and Table 2 summarizes the results. N is the number of observations used and the R^2 is adjusted. Observe that although time series of inflation are very persistent, the Fed's forecasts explain only slightly more than half of the actual variability of inflation. The Fed's forecasts of GDP growth for h > 0 are simply very poor. Private forecasters of these variables do not do any better (see Romer and Romer (2000), Tables 1 and 6) and we omit presentation of these results here.

Table 2: Accuracy of Fed's Green Book Forecasts, Estimates of (1a)-(1b)
(standard errors in parentheses)

	Inflation Forecasts				*GDP Growth Forecasts*			
h	$\alpha_0^{h,GB}$	$\alpha_1^{h,GB}$	R^2	N	$\delta_0^{h,GB}$	$\delta_1^{h,GB}$	R^2	N
0	.35 (.22)	.97 (.04)	.83	294	.80 (.30)	.89 (.08)	.53	293
1	.36 (.31)	1.00 (.07)	.72	278	.88 (.54)	.77 (.13)	.25	277
2	.32 (.36)	1.03 (.08)	.60	256	.70 (.67)	.79 (.18)	.18	255
3	.29 (.38)	1.04 (.08)	.54	239	1.05 (.93)	.62 (.27)	.08	238
4	- .13 (.41)	1.09 (.09)	.54	209	- .15 (1.07)	1.08 (.33)	.16	208
5	- .36 (.42)	1.08 (.10)	.54	150	- .70 (1.15)	1.31 (.41)	.18	149
6	- .25 (.45)	.95 (.14)	.59	90	- .80 (1.11)	1.47 (.41)	.22	89
7	- .09 (.56)	.82 (.19)	.64	59	- .26 (1.93)	1.35 (.71)	.10	58

Consider now the contribution of private forecasts, which are known to the Fed forecasters, to the forecast error of the Fed in the sense of improving the estimates of equations (1a)-(1b). We would have liked to include all three private sources and test equations of the form

(2a) $$\pi_{t,h} = \alpha_0^{h,GB} + \alpha_1^{h,GB}\pi_{t,h}^{GB} + \alpha_2^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{BLU}] + \alpha_3^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{SPF}] + \alpha_4^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{DRI}] + \varepsilon_{t,h}^{GB}$$

(2b) $$g_{t,h} = \delta_0^{h,GB} + \delta_1^{h,GB}g_{t,h}^{GB} + \delta_2^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{BLU}] + \delta_3^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{SPF}] + \delta_4^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{DRI}] + \vartheta_{t,h}^{GB}.$$

Accurate Fed forecasting implies $\alpha_j^{h,GB} = 0$, $\delta_j^{h,GB} = 0$, for j = 2, 3, 4. Due to differences in reporting dates the number of useable observations to estimate (2a)-(2b), for all three private sources jointly, is only 42 for inflation and 44 for GDP growth and hence estimates are not reliable. To increase the number of observations we examine only the joint contribution of BLU and DRI forecasts to the Green Book forecasts from 1980:1 through 1995:11. We thus estimate equations of the form

(3a) $$\pi_{t,h} = \tilde{\alpha}_0^{h,GB} + \tilde{\alpha}_1^{h,GB}\pi_{t,h}^{GB} + \tilde{\alpha}_2^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{BLU}] + \tilde{\alpha}_4^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{DRI}] + \tilde{\varepsilon}_{t,h}^{GB}$$

(3b) $$g_{t,h} = \tilde{\delta}_0^{h,GB} + \tilde{\delta}_1^{h,GB}g_{t,h}^{GB} + \tilde{\delta}_2^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{BLU}] + \tilde{\delta}_4^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{DRI}] + \tilde{\vartheta}_{t,h}^{GB}$$

for h = 0, 1, ..., 6. We exclude h = 7 since we have less than 40 observations, and Table 3 reports the results. Estimates which are significantly different from 0 at the 10% level are indicated by a (*).

Table 3: Test of Fed's Forecast Error Accuracy Using Mean BLU and DRI Forecasts
(standard errors in parentheses)

Inflation: $\pi_{t,h} = \tilde{\alpha}_0^{h,GB} + \tilde{\alpha}_1^{h,GB}\pi_{t,h}^{GB} + \tilde{\alpha}_2^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{BLU}] + \tilde{\alpha}_4^{h,GB}[\pi_{t,h}^{GB} - \pi_{t,h}^{DRI}] + \tilde{\varepsilon}_{t,h}^{GB}$

h	$\tilde{\alpha}_0^{h,GB}$	$\tilde{\alpha}_1^{h,GB}$	$\tilde{\alpha}_2^{h,GB}$	$\tilde{\alpha}_4^{h,GB}$	R^2	N
0	- .37 (.21)	1.01 (.04)	- .50* (.11)	- .39* (.11)	.93	116
1	- .16 (.31)	.94 (.06)	- .08 (.21)	.04 (.14)	.89	116
2	.21 (.31)	.88 (.06)	.46* (.20)	- .27* (.17)	.86	116
3	.63 (.29)	.78 (.06)	.68* (.24)	- .43* (.16)	.78	116
4	1.38 (.40)	.57 (.09)	.69* (.25)	- .57* (.24)	.70	112
5	1.51 (.45)	.48 (.09)	.68* (.31)	- .58* (.20)	.68	83
6	1.80 (.59)	.35 (.10)	.66* (.40)	- .69* (.27)	.52	48

GDP Growth: $g_{t,h} = \tilde{\delta}_0^{h,GB} + \tilde{\delta}_1^{h,GB}g_{t,h}^{GB} + \tilde{\delta}_2^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{BLU}] + \tilde{\delta}_4^{h,GB}[g_{t,h}^{GB} - g_{t,h}^{DRI}] + \tilde{\vartheta}_{t,h}^{GB}$

h	$\tilde{\delta}_0^{h,GB}$	$\tilde{\delta}_1^{h,GB}$	$\tilde{\delta}_2^{h,GB}$	$\tilde{\delta}_4^{h,GB}$	R^2	N
0	1.20 (.34)	.86 (.10)	.61* (.24)	- .43 (.28)	.63	127
1	.83 (1.04)	.96 (.34)	- .23 (.55)	- .62* (.38)	.23	127
2	2.32 (1.66)	.39 (.57)	.55 (.58)	- .26 (.41)	.07	127
3	3.02 (1.78)	.10 (.65)	.20 (.74)	- .10 (.30)	.02	127
4	- .22 (1.38)	1.42 (.49)	.97* (.46)	- .15 (.30)	.33	127
5	- .12 (1.32)	1.30 (.47)	1.07* (.70)	- .51 (.35)	.24	90
6	- 1.72 (2.06)	1.81 (.79)	- .02 (.44)	.35 (.24)	.42	53

Using the simple measure of R^2 it is clear the Fed's forecasts are improved by using *the known* additional information provided by the private forecasts[3]. In 10 out of the 14 equations in Table 3 there are some estimates of $(\alpha_j^{h,GB}, \delta_j^{h,GB})$ for $j = 2, 4$ which are significantly different from zero. Indeed, for any combination of private forecasters there are horizons for which private forecasts significantly improve the forecast error of the Fed forecasts in (3a)-(3b). Hence, the known private forecasts are correlated with the Fed's forecast error. This conclusion raises natural questions about the accuracy of the Fed forecasts. It also leads one to question if there is any sense in which one may presume the Fed's staff uses a "correct" model to forecast inflation and GDP growth.

4.2 *Errors in Forecasting the Effect of Monetary Policy: Structural or Reduced Form Analysis?*

We aim to test if the Fed and private forecasters use correctly specified models of inflation and GDP growth. Had our economic environment remained stochastically stationary with a fixed joint density of events over time and had the Fed used a constant and credible policy rule, forecasters would learn the correct models from past data. Absent such an environment we examine how well do forecasters adapt to these changes. To do that we study the structure of forecast errors over the three monetary regimes after 1980. The question is what econometric model we should use.

Since we study forecasters' use of incorrect models, *a structural analysis* requires a formulation of a structural model of misspecification! Keeping in mind that agents do not learn the true structure since it changes over time with technology and institutions, misspecification varies over time: errors in 1965:11-1979:8 are different from errors in 1979:9 - 1987:8. Since unlimited misspecified models are possible, the task at hand is daunting. But, we face a deeper problem. A structural analysis requires us, the researchers, to have some prior guess or knowledge of the structure of misspecification by forecasters at each era. But if we had such knowledge, the forecasters would know it as well and the misspecification would not occur. With this in mind we observe that a retrospective structural analysis is a study which can merely reveal *why forecasters made one error or another*. But since the causes of forecasting errors vary over time, it would be rare to find persistent error with a universal lesson. Moreover, without a "theory" to motivate modeling of the errors, the data would be compatible with many different models. Such circularity suggests that a structural approach to a retrospective study of

[3] Romer and Romer (2000) report similar tests for inflation in their Table 4 and for GDP growth in their Table 6. Their Table 4 reports that 10 out of 11 equations contain estimates of similar parameters which are significantly different from zero and their Table 6 reports 7 out of 20 such equations. Our conclusions are also compatible with Sims (2002).

the accuracy of forecasting is not the right approach.

The alternative methodology is a *reduced form approach*, which focuses on predictability of forecast errors by present date observables. The key variables on which we focus are the monetary policy regimes in place when the forecasts were made. Since a change in a monetary regime entails a change in the effective monetary rule, we ask how changes in the rule affect the forecast error of the Fed staff and private forecasters. Does the Fed forecasters assess correctly the effect of monetary policy on future GDP growth and inflation? Fed forecasters could have failed to assess the *direct* effect of interest rates on future inflation and GDP growth. But their models could incorrectly specify the effect of other variables and their interaction with monetary rules, resulting in an incorrect assessment of the *indirect* effect of interest rates on future inflation and growth. Since we answer the questions with a reduced form model, an interpretation of the results must be understood to be in the reduced form sense. But then, this is all that matters. We do not explain why forecasters failed to correctly assess the direct effect of interest rates on inflation and growth. We only ask if they correctly assessed the total effect, direct and indirect, of the monetary rule and such a failure may have been caused by many different misspecifications. From a practical point of view all that matters is the accuracy of the forecasts and the variables which explain the failure to forecast accurately. To accomplish this we have defined two sets of variables:

(i) ***Time Period***. We divided the sample period into three sub-periods. Sub-period 1 is the pre-Volcker period 1965:11-1979:8, sub-period 2 is the Volcker era 1979:9-1987:8 and sub-period 3 is part of the Greenspan era starting in 1987:9 and ending in 1995:11. Hence we define

$$1_j(t) = \begin{cases} 1 & \text{if} \quad t \in \text{sub-period } j \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Although these time periods reflect differences of policy rule in place and in leadership of the Fed, they also reflect changes in society, technology and economic institutions. Since we estimate the forecast error functions associated with each period, *they reflect all changes in the environment*.

(ii) ***Known measures of monetary policy***. We utilize two measures. The first is the Federal Funds Rate at the end of t-1, denoted by FF_{t-1} and known at month t, the forecasting date. To define the second variable we first identify the usual monetary regimes: the regime of *tightening* monetary policy defined by *rising* funds rates and the regime of *loosening* monetary policy defined by *falling* funds rates. We define CFF_{t-1} by first identifying the monetary regime in place at date t-1. Then, CFF_{t-1} *is the cumulative change of the Fed Funds rate from the start of the present monetary regime until the end of the month at t-1.* Note that we use only variables dated t-1 to ensure that regardless of the month when a forecaster makes his forecasts, he knows the exact state of the monetary variables defined here. He only needs to predict their impact.

In addition to monetary policy variables, we also include, as an explanatory variable, *the forecast itself*. This variable measures the degree to which a forecaster incorrectly assessed the impact of *all other variables in the economic environment* on future rates of inflation and GDP growth. Given these we let k = GB, BLU, SPF and DRI. We then estimate the following model which extends the earlier model (1a)-(1b) as follows:

(5a) $$\pi_{t,h} - \pi^{k}_{t,h} = \sum_{j=1}^{3} 1_j(t)\left[\hat{\alpha}^{h,k}_{0,j} + \hat{\alpha}^{h,k}_{1,j}\pi^{k}_{t,h} + \hat{\alpha}^{h,k}_{2,j}FF_{t-1} + \hat{\alpha}^{h,k}_{3,j}CFF_{t-1}\right] + \hat{\varepsilon}^{k}_{t,h}$$

(5b) $$g_{t,h} - g^{k}_{t,h} = \sum_{j=1}^{3} 1_j(t)\left[\hat{\delta}^{h,k}_{0,j} + \hat{\delta}^{h,k}_{1,j}g^{k}_{t,h} + \hat{\delta}^{h,k}_{2,j}FF_{t-1} + \hat{\delta}^{h,k}_{3,j}CFF_{t-1}\right] + \hat{\upsilon}^{k}_{t,h}.$$

For each k and for each horizon we estimate 12 parameters for every equation. We ignore the horizon h = 7 since the number of observations is too small.

4.2.1 *What Does Theory Predict Agents Do in This Environment?*

Our null hypothesis is the *perfect knowledge* hypothesis under which $\hat{\alpha}^{h,k}_{ij} = \hat{\delta}^{h,k}_{i,j} = 0$ *for all i, j, h, k.* In all tables below we provide t-tests for each parameter and Wald tests for the composite hypothesis. But what does theory say we should expect to find and how should we interpret parameters which are significantly different from zero? When forecasters know the true probability model and use all available information, (5a)-(5b) are pure noise. If they are not noise, we view them as error functions of forecasters who did not know the true structure and could not learn it fast enough. Our agents are perfectly rational and given the changing economic conditions, they learn the best they can, given the limited data they obtain within each sub-period. When the environment keeps changing, the agent's models never converge hence they never attain perfect knowledge.

The theory of Rational Belief (in short, RB) due to Kurz (1994), (1997) assumes the economic environment changes but agents have a large bank of past data from which they learn the long term empirical distribution of observables. This known distribution is the discipline on their beliefs. This distribution implies a unique probability measure on infinite sequences of observables which is stationary and known to all agents. *This is the non-judgmental probability of future events*. A belief is said to be an RB if it is a probability model of the economy which, if simulated, *reproduces the known empirical distribution*. Since there are many models which reproduce the same empirical distribution, the RB theory predicts agents would use diverse subjective models to forecast inflation and GDP growth. Since the models are not based on perfect knowledge, they exhibit significant error functions. The estimated error functions are significant since agents could not, at the time, learn the structure of each regime fast enough. In that case (5a)-(5b) is a retrospective recording of how well they adapted to the changing conditions. But what about the long run?

The RB rationality principle implies that over long enough time forecasters learn *not to persist with the same type of error* but will not be able to avoid different errors. The theory thus predicts that the error functions in different regimes will be different. However, over sufficiently long time period subjective models are disciplined by the known empirical distribution. Hence, the RB theory predicts that for sufficiently long time the *average of the forecast functions converge to zero*. This is intuitively clear: forecasters know it is difficult to forecast GDP growth three quarters from today and exhibit wide diversity of such forecasts. But they also know it is easier to forecast average GDP growth over the next five years and will exhibit smaller diversity in making such forecasts.

Is the period 1965-1995 long enough for long term averaging of the forecast errors? To test this we estimated (5a)-(5b) for the entire period, disregarding the three regimes in the sub-periods. Apart from some important exceptions we discuss later, the results confirm the RB prediction. However, this conclusion is not helpful for policy decisions which require accurate forecasting of 1-2 years. We thus expect patterns of forecasting errors to be *quantitatively* different across the three regimes although *qualitative similarity* across sub-periods is possible and is, in fact, exhibited in the data.

Since our estimates evaluate accuracy of forecasting models within each sub-period, a correct reading of the results in the tables below requires us to focus on what a forecaster "got wrong" or assessed incorrectly. We therefore comment now on the interpretation of the parameters.

4.2.2 *Interpreting the Estimated Parameters of Forecast Error Functions*

Inflation Forecasts. We use the forecasts $(\pi^k_{t,h}, g^k_{t,h})$ as explanatory variables to test the effect of *all other information* available at t, incorporated in $(\pi^k_{t,h}, g^k_{t,h})$ but not explicit in (5a)-(5b). We thus suggest that the parameters $(\hat{\alpha}^{h,k}_{0,j}, \hat{\delta}^{h,k}_{0,j})$ and $(\hat{\alpha}^{h,k}_{1,j}, \hat{\delta}^{h,k}_{1,j})$ tell us something about the subjective model used by a forecaster. The typical results are $(\hat{\alpha}_{0,j}^{h,k} > 0, \hat{\delta}_{0,j}^{h,k} > 0)$ and $(\hat{\alpha}_{1,j}^{h,k} < 0, \hat{\delta}_{1,j}^{h,k} < 0)$, and we illustrate the interpretation of the parameters by explaining the meaning of this pattern. We discuss here in detail the interpretation and then the results for inflation forecasts as in (5a). Once the details are worked out, the reader can apply similar reasoning to the results of the GDP growth in (5b) and these are reported and discussed in Section 4.4:

(i) <u>A positive estimate</u> of $\hat{\alpha}_{0,j}^{h,k}$ means that in period j and for horizon h the *intercept* of the forecast function was too low hence forecaster k generated forecasts that were, *on average,* lower than the mean inflation realized later. That is, forecasts *underestimated* future inflation. $\hat{\alpha}_{0,j}^{h,k} < 0$ means the forecasts *overestimated* future inflation.

(ii) <u>A negative estimate of</u> $\hat{\alpha}_{1,j}^{h,k}$ means the regression coefficient of forecasts on realizations was less than 1: a 1% change in forecasted inflation was associated, on average, with a realized rate which was smaller than 1%. Forecasts implied a higher variance than was in the data.

(iii) A combination ($\hat{\alpha}_{0,j}^{h,k} > 0$, $\hat{\alpha}_{1,j}^{h,k} < 0$) implies *underestimation* of inflation when inflation forecasts are low and *overestimation* of inflation when inflation forecasts are high. Hence the model generates forecasts with too high variance: too low when low and too high when high!

Monetary Policy Variables. We now interpret $\hat{\alpha}_{2,j}^{h,k}$, the parameters of FF_{t-1}. It measures how well forecaster k assessed the impact of changes in the funds rate on future inflation rates over horizon h in monetary regime j. Since the capital\output ratio is almost constant within each policy regime, the real rate of return on capital is constant. Hence the nominal rate FF_{t-1} is treated as a cause of change in subsequent inflation and GDP growth rates. Since a high funds rate is associated with *lower* future inflation, $\hat{\alpha}_{2,j}^{h,k} > 0$ says that a high funds rate leads to a positive forecast error hence to forecasted inflation which, on average, is lower than realized inflation. Too low forecasted inflation means, on average, an *overestimation*[4] of the impact of high funds rate on reducing future inflation. Similarly, a <u>*negative*</u> parameter means the forecaster did not assess correctly the effect of monetary policy and, on average, *underestimated* the ability of the Fed to reduce future inflation by raising rates.

[4] The terminology we adopt here may be confusing a bit. It is based on the fact that a high federal funds rate aims to *reduce* inflation. The fact is that a positive coefficient of either FF_{t-1} or CFF_{t-1} means that a positive value of FF_{t-1} or CFF_{t-1} leads to expected values which are *smaller* than conditional mean realizations and we would often think of this as an *underestimate* by a forecaster. Here we say that a forecaster overestimated the ability of funds rate or the cumulative funds rate to lower inflation. Since the aim of policy is to reduce the inflation rate, if the expected rate is lower on average than the realized rates we interpret this to mean the forecaster overestimated the power of policy to attain its goal.

What does CFF_{t-1} measures? A 5% Fed Funds rate after a period of rising interest rates (i.e., $CFF_{t-1} > 0$ and large) may be different from a 5% Fed Funds rate at the start of new cycle of rising rates (i.e., $CFF_{t-1} > 0$ and small). The long run average of CFF_{t-1} is zero, it moves away from zero as the duration of a regime increases and it reaches its largest values just before the regime changes. $\hat{\alpha}_{3,j}^{h,k} > 0$ means that cumulative rising rates generate overshooting forecasts or *positive* forecast errors. Such forecaster *overestimates* the cumulative effects of monetary policy on future inflation: he predicted too big *increases* of inflation rates in response to cumulative declines of funds rate and predicted too large a *decline* of inflation in response to cumulative increase in rates.

4.3 *Testing the Accuracy of Inflation Forecasting (Equation (5a)*)

Results for inflation forecasting (5a) are reported in Tables 4A-4D and for GDP growth (5b) are reported later in Tables 7A - 7D. As before, the symbol (*) means that the parameter is statistically significantly different from zero *at least* at a 10% confidence level.

For each h from 0 to 6 parameters are arranged in four groups. Reading down each column we report first the three constants, second- the three parameters of the forecasts, third- the three parameters of FF_{t-1} and finally the three parameters of CFF_{t-1}. All R^2 are adjusted for degrees of freedom and (*) indicates statistical significance of a t test of at least 10%. The χ^2 statistic is for a composite Wald test of the perfect knowledge null hypothesis $\hat{\alpha}_{i,j}^{h,k} = 0$ for all i, j in that column.

4.3.1 *Assessing The Fed's Inflation Forecast Accuracy*

We start with a detailed examination of the results for the Fed's forecasts, reported in Table 4A.

4.3.1.a *The Fed's forecasts during the 1965-1979 monetary regime*

(i) $\hat{\alpha}_{0,1}^{h,GB} > 0$ for long horizon forecasts. Hence Fed forecasts underestimated future inflation: the average realized inflation was significantly higher than the average Fed's forecasts.

(ii) $\hat{\alpha}_{1,1}^{h,GB} < 0$ for all horizons. The regression coefficient of forecasts on realizations was less than 1: for every percentage point of forecasted inflation, realized inflation was on average proportionally lower. Combination (i)-(ii) implies the Fed forecasted too low inflation when inflation forecasts were low and too high inflation when forecasts were high.

(iii) $\hat{\alpha}_{2,1}^{h,GB} > 0$ for short horizons. The Fed overestimated the impact of higher interest rates on reducing inflation over short horizons. The mean decline in inflation in response to higher funds rate was, on average, not as large as the Fed expected.

(iv) $\hat{\alpha}_{3,1}^{h,GB} > 0$ for long horizons. The Fed overestimated the cumulative effects of monetary policy on inflation over long horizons. The Fed forecasted too low inflation in response to cumulative rising rates and it forecasted too high inflation in response to cumulative decreasing rates.

In sum, during this era the Fed forecasters underestimated future inflation when inflation forecasts were low and overestimated future inflation when inflation forecasts were high. In addition, the Fed overestimated the impact of monetary policy on inflation over all horizons.

Table 4A: Estimated Fed's Inflation Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4	5	6
$\hat{\alpha}_{0,1}^{h,GB}$	-.03	-.02	1.46	3.01*	6.06*	11.11*	20.30*
	(.43)	(.70)	(1.14)	(1.17)	(1.18)	(1.19)	(1.30)
$\hat{\alpha}_{0,2}^{h,GB}$	-.02	.70*	.11	.28	.15	.47	1.21
	(.61)	(.41)	(.51)	(.48)	(.62)	(.58)	(1.02)
$\hat{\alpha}_{0,3}^{h,GB}$	.02	-.17	-.21	.28	.13	.36	.65
	(.30)	(.36)	(.35)	(.42)	(.38)	(.43)	(.43)
$\hat{\alpha}_{1,1}^{h,GB}$	-.29*	-.34*	-.41*	-.46*	-.79*	-1.45*	-.54*
	(.06)	(.11)	(.15)	(.14)	(.22)	(.21)	(.22)
$\hat{\alpha}_{1,2}^{h,GB}$	-.10	.05	.09	.09	.15	-.03	-.48*
	(.10)	(.06)	(.08)	(.08)	(.14)	(.14)	(.25)
$\hat{\alpha}_{1,3}^{h,GB}$	-.39*	-.42*	-.03	-.03	.20	-.12	-.39*
	(.12)	(.14)	(.17)	(.17)	(.16)	(.24)	(.19)
$\hat{\alpha}_{2,1}^{h,GB}$	.27*	.42*	.32*	.13	-.04	- .10	-1.76*
	(.08)	(.14)	(.19)	(.19)	(.13)	(.07)	(.08)
$\hat{\alpha}_{2,2}^{h,GB}$	.02	-.14*	-.11*	-.14*	-.17*	-.14*	-.01
	(.08)	(.04)	(.04)	(.04)	(.07)	(.07)	(.10)
$\hat{\alpha}_{2,3}^{h,GB}$	.15*	.20*	-.01	-.08	-.17*	-.06	-.01
	(.07)	(.09)	(.10)	(.10)	(.10)	(.10)	(.11)
$\hat{\alpha}_{3,1}^{h,GB}$	-.04	-.03	.08	.27*	.49*	.60*	.78*
	(.05)	(.07)	(.10)	(.09)	(.10)	(.06)	(.05)
$\hat{\alpha}_{3,2}^{h,GB}$	-.01	.03	.02	.07*	.02	.05	.11*
	(.05)	(.02)	(.03)	(.04)	(.05)	(.04)	(.06)
$\hat{\alpha}_{3,3}^{h,GB}$	-.03	-.04	.01	.04	.03	.01	.01
	(.03)	(.04)	(.04)	(.04)	(.03)	(.03)	(.04)
R^2	.28	.42	.43	.51	.68	.81	.71
N	294	278	256	239	209	150	90
Chi Square value for test: $\hat{\alpha}_{i,j}^{h,GB} = 0, \forall i,j$	102.3	123.9	113.9	118.1	131.5	334.8	584.2

4.3.1.b *The Fed's forecasts during the Volcker monetary regime*

(i) Both $(\hat{\alpha}_{0,2}^{h,GB}, \hat{\alpha}_{1,2}^{h,GB})$ are not different from 0. The Fed forecasters assessed mean inflation correctly and the regression coefficient of forecasts on realizations was not significantly different from 1.

(ii) In contrast with the earlier pattern during this era $\hat{\alpha}_{2,2}^{h,GB} < 0$: the Fed underestimated the impact of higher interest rates on lowering inflation over all horizons.

(iii) $\hat{\alpha}_{3,2}^{h,GB} > 0$ for some long horizons: the Fed overestimated cumulative effects of monetary policy on long horizon inflation. This secondary effect is small.

In sum, Table 4A shows that during this era the Fed forecasters underestimated the impact of tight monetary policy on inflation over all horizons. However, we also find small overestimation of cumulative effects of tight money on long horizon inflation rates. We reconcile the two conclusions by noting that result (ii) reflects an underestimation of the impact of tight money policy on inflation in all horizons. Result (iii) expresses the fact that the underestimation in (ii) tended to weaken after a *long period* of rising rates. The decrease in bias is small.

4.3.1.c ***The Fed's forecasts during the Greenspan monetary regime***

(i) During the Greenspan era the Fed forecasters assessed mean inflation correctly (i.e. $\hat{\alpha}_{0,2}^{h,GB} = 0$) .

(ii) $\hat{\alpha}_{1,3}^{h,GB} < 0$ for short and very long (6 quarters) horizons: their inflation forecasts were too high.

(ii) $\hat{\alpha}_{2,3}^{h,GB} > 0$ for h = 0, 1: the Fed overestimated the impact of higher interest rates on inflation over short horizons of up to one quarter.

In sum, during this era the Fed forecasted too high inflation rates for some horizons. They also overestimated the ability of higher interest rates to lower inflation over short horizons. This means that Fed forecasters incorrectly expected policy to impact the economy earlier than it did.

Summary of Results for the Fed Inflation Forecasting Error Functions

(1) Table 4A exhibits many significant parameters which are different across the three periods.

(2) The R^2 of the error functions are surprisingly very high, rising with the length of the horizon. The estimated error functions account for over 65% of all long horizon inflation forecast errors!

(3) During the first and third periods $\hat{\alpha}_{1,1}^{h,k} < 0$, $\hat{\alpha}_{1,3}^{h,k} < 0$ for all forecast horizons. These imply the Fed forecasted too low inflation when inflation forecasts were low and too high inflation when high.

(4) The Fed never assessed the correct effect of rates on inflation. In the first and third periods the Fed overestimated and in the second it underestimated the stabilization effects of interest rates.

(5) The Chi Square values of the Wald composite test $\hat{\alpha}_{i,j}^{h,GB} = 0 , \forall i,j$ are very large for all h. With 12 degrees of freedom the 10% critical value of a Chi Square test is 18.5 and the 1% value is 26.2 hence the perfect knowledge null hypothesis $\hat{\alpha}_{i,j}^{h,GB} = 0$ all i and j is rejected for all h.

4.3.2 ***Assessing the Accuracy of Private Inflation Forecasting Models.***

A discussion of the private forecasting models utilizes the same interpretation of parameters as the one of the Fed forecasting model. Considering Tables 4B - 4D together note that since DRI data cover the same period as the Fed, Table 4B covers all monetary regimes as Table 4A. SPF forecasts cover only four future quarters hence the forecasting horizon in Table 4C is shorter than in Table 4A. Finally, since BLU data starts in 1980, Table 4D covers only the second and third monetary eras.

Before commenting on specific forecasters we make an important observation on Tables 4A- 4D. Although there are significant differences which we discuss in detail later, the sign pattern of estimated parameters and the pattern of statistical significance of coefficients *are surprisingly similar in all four tables*. We thus observe that the forecast errors of the Fed and those of the private forecasters have very similar characteristics.

Table 4B: Estimated DRI's Inflation Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4	5	6	7
$\hat{\alpha}_{0,1}^{h,DRI}$	.14	.83	2.11*	3.37*	4.90*	6.70*	8.02*	9.12*
	(.66)	(.70)	(1.06)	(1.28)	(1.13)	(1.09)	(1.39)	(1.40)
$\hat{\alpha}_{0,2}^{h,DRI}$	-.19	.74*	.80*	1.69*	2.15*	2.37*	2.10*	2.32*
	(.46)	(.39)	(.36)	(.54)	(.60)	(.42)	(.40)	(.50)
$\hat{\alpha}_{0,3}^{h,DRI}$	.19	1.12	.58	1.12	1.91*	1.31	1.84*	2.76*
	(.77)	(1.15)	(.56)	(.82)	(.67)	(.94)	(1.11)	(1.28)
$\hat{\alpha}_{1,1}^{h,DRI}$	- .49*	- .84*	-.88*	-.77*	-.81*	-.82*	-.72*	-.55*
	(.15)	(.13)	(.18)	(.22)	(.22)	(.22)	(.25)	(.22)
$\hat{\alpha}_{1,2}^{h,DRI}$	-.16*	.02	.18	.03	.03	-.00	-.14	-.35*
	(.08)	(.13)	(.13)	(.11)	(.16)	(.16)	(.14)	(.12)
$\hat{\alpha}_{1,3}^{h,DRI}$	-.65*	-.95*	-.68*	-.66*	-.66*	-.45*	-.53*	-.59*
	(.13)	(.22)	(.17)	(.16)	(.20)	(.20)	(.16)	(.13)
$\hat{\alpha}_{2,1}^{h,DRI}$	.53*	.76*	.61*	.33	.14	- .11	- .33*	- .58*
	(.17)	(.14)	(.24)	(.21)	(.15)	(.10)	(.13)	(.15)
$\hat{\alpha}_{2,2}^{h,DRI}$	.05	-.14*	-.24*	-.28*	-.36*	-.39*	-.31*	-.24*
	(.05)	(.07)	(.06)	(.08)	(.11)	(.09)	(.08)	(.08)
$\hat{\alpha}_{2,3}^{h,DRI}$	.28*	.30*	.24*	.15	.04	.02	-.05	-.14
	(.10)	(.09)	(.09)	(.09)	(.11)	(.12)	(.12)	(.13)
$\hat{\alpha}_{3,1}^{h,DRI}$	-.11	-.07	.03	.23*	.43*	.55*	.47*	.42*
	(.08)	(.07)	(.11)	(.09)	(.08)	(.11)	(.14)	(.12)
$\hat{\alpha}_{3,2}^{h,DRI}$	.03	.02	.03	.15*	.16*	.16*	.13*	.11*
	(.03)	(.05)	(.04)	(.04)	(.06)	(.05)	(.03)	(.04)
$\hat{\alpha}_{3,3}^{h,DRI}$	-.06	-.01	-.00	.05	.12*	.12	.14*	.22*
	(.06)	(.07)	(.06)	(.07)	(.06)	(.08)	(.09)	(.10)
R^2	.37	.55	.52	.59	.71	.78	.77	.80
N	248	248	248	248	248	248	248	246
Chi Square value for test: $\hat{\alpha}_{i,j}^{h,DRI}=0, \forall i,j$	80.6	198.6	148.6	141.4	201.4	500.7	663.9	718.1

4.3.2.a *DRI Inflation Forecasting Error*

In Table 4B $\hat{\alpha}_{0,j}^{h,DRI} > 0$, all j and most $\hat{\alpha}_{1,j}^{h,DRI}$ are negative: DRI's forecasts were too low when inflation forecasts were low and too high when forecasts were high. The effects of monetary policy on DRI's forecast error is qualitatively similar to the Fed's in Table 4A: it overestimated the effect of stabilization for short horizons and underestimated it for long horizons. The R^2 are extremely high, rising with the forecast horizon. Finally, the Chi Square values of the test $\hat{\alpha}_{i,j}^{h,DRI} = 0, \forall i, \forall j$ are very large. With 12 degrees of freedom the 10% critical value of a Chi Square test is 18.5 and the 1% value is 26.2 hence the perfect knowledge null hypothesis is rejected by the data for all h. We now turn to the SPF equation reported in Table 4C and make several observations.

4.3.2.b *SPF Inflation Forecasting Error*

During the 1965-1979 monetary regime:

(i) $\hat{\alpha}_{0,1}^{h,SPF} > 0$ for long horizons (3 - 4 quarters): SPF forecasts underestimated future inflation rates.

(ii) $\hat{\alpha}_{1,1}^{h,SPF} < 0$ for all horizons and increasing with the horizon.

(iii) $\hat{\alpha}_{2,1}^{h,SPF} > 0$ for short horizons: SPF overestimated effects of high interest rates on lower inflation.

(iv) $\hat{\alpha}_{3,1}^{h,SPF} > 0$ for long horizons: SPF overestimated the cumulative effects of monetary policy.

Table 4C: Estimated Median SPF Inflation Forecast Error Functions
(standard errors in parentheses)

Horizon →	0	1	2	3	4
$\hat{\alpha}_{0,1}^{h,SPF}$	.64 (.86)	.33 (.85)	1.55 (1.20)	3.34* (1.22)	3.90* (1.36)
$\hat{\alpha}_{0,2}^{h,SPF}$	-.52 (.49)	-.17 (.60)	-.88 (.74)	-.70 (.90)	1.42* (.68)
$\hat{\alpha}_{0,3}^{h,SPF}$	-.39 (.41)	-.53 (.47)	-.85* (.50)	-.74 (.52)	-.65 (.54)
$\hat{\alpha}_{1,1}^{h,SPF}$	-.32* (.10)	-.47* (.13)	-.47* (.15)	-.59* (.14)	-.65* (.19)
$\hat{\alpha}_{1,2}^{h,SPF}$	-.01 (.17)	.27* (.14)	.38 (.25)	.21 (.20)	.18 (.24)
$\hat{\alpha}_{1,3}^{h,SPF}$	-.25* (.13)	-.29* (.16)	-.37* (.19)	-.35* (.20)	-.41* (.24)
$\hat{\alpha}_{2,1}^{h,SPF}$	.29* (.15)	.49* (.14)	.33 (.20)	.17 (.21)	.16 (.15)
$\hat{\alpha}_{2,2}^{h,SPF}$	.01 (.08)	-.20* (.08)	-.22* (.11)	-.18* (.11)	-.38* (.12)
$\hat{\alpha}_{2,3}^{h,SPF}$	.13 (.09)	.16* (.10)	.23* (.10)	.19* (.11)	.20* (.09)
$\hat{\alpha}_{3,1}^{h,SPF}$	-.01 (.08)	.02 (.08)	.13 (.11)	.26* (.10)	.43* (.10)
$\hat{\alpha}_{3,2}^{h,SPF}$	.04 (.05)	.09* (.05)	.00 (.07)	.22* (.07)	.20* (.10)
$\hat{\alpha}_{3,3}^{h,SPF}$	-.05 (.04)	-.05 (.04)	-.08* (.05)	-.05 (.05)	-.06 (.04)
R^2	.30	.54	.54	.58	.67
N	129	128	127	126	120
Chi Square value for test: $\hat{\alpha}_{i,j}^{h,SPF} = 0, \forall i,j$	67.8	117.1	143.8	168.9	261.8

During the Volcker Monetary Regime:

(i) SPF underestimated the effect of higher interest rates on reducing inflation over all horizons.

(ii) SPF overestimated the cumulative effects of monetary policy on reducing inflation in longer forecasting horizons (3 - 4 quarters).

During the Greenspan monetary regime:

(i) SPF were strongly biased upward for all forecasting horizons.

(ii) SPF overestimated the effect of higher interest rates on reducing inflation over all horizons. This is reflected in $\hat{\alpha}_{2,3}^{h,SPF} > 0$ for all horizons. Recall that for the Fed this systematic forecasting mistake is exhibited only for short horizons while for the median SPF it is present for all horizons. This is the main difference between the Fed's error and the median SPF error.

The Chi Square values of the composite Wald test $\hat{\alpha}_{i,j}^{h,SPF} = 0, \forall i, j$ are very large. With 12 degrees of freedom the 10% critical value of a Chi Square test is 18.5 and the 1% value is 26.2 hence the composite perfect knowledge hypothesis $\hat{\alpha}_{i,j}^{h,SPF} = 0, \forall i, j$ is rejected by the data for all h.

Table 4D: Estimated Mean BLU Inflation Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4	5	6
$\hat{\alpha}^{h,BLU}_{0,2}$	-.38	-.39	-.99*	-.96	-.24	.22	2.17*
	(.37)	(.46)	(.53)	(.65)	(.64)	(.79)	(1.04)
$\hat{\alpha}^{h,BLU}_{0,3}$	-.38	-.56	-.77*	-.76*	-.70	-.73	-.46
	(.37)	(.40)	(.41)	(.45)	(.49)	(.48)	(.51)
$\hat{\alpha}^{h,BLU}_{1,2}$	.12	.50*	.68*	.60*	.16	-.06	-1.07*
	(.10)	(.14)	(.18)	(.20)	(.20)	(.25)	(.38)
$\hat{\alpha}^{h,k}_{1,3}$	-.23*	-.28*	-.30*	-.24	-.32*	-.31*	-.23
	(.13)	(.15)	(.17)	(.19)	(.20)	(.19)	(.17)
$\hat{\alpha}^{h,BLU}_{2,2}$	-.08	-.32*	-.39*	-.38*	-.25*	-.19*	.16
	(.06)	(.09)	(.08)	(.09)	(.10)	(.10)	(.15)
$\hat{\alpha}^{h,BLU}_{2,3}$	.11	.15*	.18*	.15*	.16*	.12*	.01
	(.09)	(.09)	(.09)	(.09)	(.08)	(.07)	(.08)
$\hat{\alpha}^{h,BLU}_{3,2}$	.05	.05*	.09*	.16*	.00	-.00	.01
	(.04)	(.03)	(.04)	(.05)	(.05)	(.06)	(.05)
$\hat{\alpha}^{h,BLU}_{3,3}$	-.04	-.05	-.05	-.03	-.04	-.01	.01
	(.03)	(.04)	(.04)	(.04)	(.04)	(.03)	(.03)
R^2	.05	.26	.36	.38	.44	.52	.65
N	252	249	246	243	235	174	114
Chi Square value for test: $\hat{\alpha}^{h,BLU}_{i,j} = 0, \forall i,j$	65.5	104.2	284.0	180.5	220.8	577.0	672.0

4.3.2.c ***BLU Inflation Forecasting Error***

Restricting attention to the two monetary regimes after 1980, Table 4D shows the differences between the pattern of forecast errors of the consensus BLU and those of the Fed are as follows:

(i) $\hat{\alpha}^{h,BLU}_{0,3} < 0$ for moderate horizons: the BLU consensus inflation forecasts were too high relative to the mean realized inflation rate during the Greenspan era. This pattern does not hold for the Fed.

(ii) $\hat{\alpha}^{h,k}_{1,2} > 0$ for h = 1, 2, 3 during the Volcker era and *this is unusual*: for all models combined, out of 74 estimated coefficients $\hat{\alpha}^{h,k}_{1,j}$ only four are positive and significantly different from zero.

(iii) $\hat{\alpha}^{h,k}_{2,3} > 0$ over all horizons during the Greenspan era. Similar to the median SPF, the consensus BLU overestimated the effect of higher interest rates on reducing inflation. This is an important difference between the Fed's error and the forecast errors of SPF and BLU.

The Chi Square values of the Wald tests $\hat{\alpha}^{h,BLU}_{i,j} = 0, \forall i,j$ are very large for every h. With 8 degrees of freedom the critical value for a 10% confidence is 13.3 and for a 1% confidence is 20.1 hence, again, the perfect knowledge null hypothesis is rejected by the data.

4.3.3 ***Comparing Error Functions***

The χ^2 tests in tables 4A-4D universally reject the perfect knowledge hypothesis. The correlation coefficient R^2 is very high for all forecasters and for all horizons, revealing the degree to which market dynamics moves too fast for forecasters to learn the correct structure. Agents have subjective models which exhibit systematic errors within each period and these errors are different across monetary regimes. However, we have also arrived at a surprising conclusion.

Although we find great heterogeneity of forecasting models, we also find similar *qualitative* patterns of forecast error functions, reflected in the signs of parameters across participants. Such qualitative similarity suggests similarity in concepts and ideas underlying the forecasting models of market participants *even when these models are wrong,* particularly with respect to the effect of monetary policy. It suggests a high degree of correlation among beliefs and subjective models of agents, confirming earlier observations related to Tables 1A-1B. Correlation among forecast error functions of agents is a powerful force generating market volatility which Kurz (1974),(1997) calls "Endogenous Uncertainty" (for recent work see Kurz, Jin and Motolese (2005a), (2005b)).

The qualitative similarity is noteworthy for an additional reason. Agency considerations were used by some (e.g. Scharfstein and Stein (1990), and Ehrbeck and Waldman (1996)) to suggest that private forecasts may be suspect due to a possible motive to distort their published forecasts. The similarity of the observed pattern of forecast error functions provides additional support to the view that no individual bias is present although a herding incentive may be present in model structures[5].

Appearance of qualitative similarity should not obscure the differences among forecast errors. To explore them we include in this discussion the results on forecasting GDP growth reported in Tables 7A-7D[6]. There are two noted differences between the Fed and private forecast errors. The first is the difference between the Fed and private forecasters' assessment of the variables under

Table 5: Proportion of Positive Parameter Estimates

(Tables 4A-4D on Inflation and 7A-7D on GDP Growth Combined)

Parameter Group	Fed	DRI	SPF	BLU
$(\hat{\alpha}^{h,k}_{0,j}, \hat{\delta}^{h,k}_{0,j})$	0.79	0.91	0.57	0.47
$(\hat{\alpha}^{h,k}_{1,j}, \hat{\delta}^{h,k}_{1,j})$	0.24	0.20	0.21	0.23
$(\hat{\alpha}^{h,k}_{2,j}, \hat{\delta}^{h,k}_{2,j})$	0.26	0.36	0.29	0.40
$(\hat{\alpha}^{h,k}_{3,j}, \hat{\delta}^{h,k}_{3,j})$	0.64	0.68	0.57	0.47

consideration. Table 5 examines the proportion of positive parameter estimates and exhibits the general pattern of the sign distribution which we have noted earlier, namely that the majority of

[5] Our results are also compatible with the work of authors who try, in retrospect, to construct improved forecasting models by using the estimated forecast errors (see, for example Shaffer (1998) ,(2003)).

[6] Tables 7A-7D are reported in section 4.4 on forecasting GDP growth. Here we merely count the total number of positive point estimates and statistically significant estimates in order to draw a general idea on qualitative differences between the Fed and private forecasters. A detailed discussion of the results for GDP growth reported in Tables 7A-7D is provided later.

parameters satisfy $(\hat{\alpha}_{0,j}^{h,k}>0, \hat{\delta}_{0,j}^{h,k}>0)$, $(\hat{\alpha}_{3,j}^{h,k}>0, \hat{\delta}_{3,j}^{h,k}>0)$, $(\hat{\alpha}_{1,j}^{h,k}<0, \hat{\delta}_{1,j}^{h,k}<0)$, $(\hat{\alpha}_{2,j}^{h,k}<0, \hat{\delta}_{2,j}^{h,k}<0)$. Table 5 reveal large quantitative differences among forecasters which lead to major differences in forecasts.

Next, consider the pattern of parameters which are significantly different from zero. Table 6 distinguishes between parameters associated with monetary policy and those associated with other variables. It shows the Fed and private forecasters made similar errors in assessing the effect of monetary policy: 44% for the Fed and 49% for private forecasters. The differences are in parameters $(\hat{\alpha}_{0,j}^{h,k}, \hat{\delta}_{0,j}^{h,k})$ and $(\hat{\alpha}_{1,j}^{h,k}, \hat{\delta}_{1,j}^{h,k})$. The proportion of significant non monetary policy parameters is only 36% for the Fed but 64% for private forecasters. Thus, in forecasting future inflation and GDP growth the relatively larger error of the Fed's model is found in its failure to assess correctly the impact of monetary policy on these variables. Private forecasters erred less in assessing the effect of monetary policy relative to other observable macro-economic variables.

Table 6: Proportion of Statistically Significant Parameter Estimates
(Tables 4A-4D and 7A-7D Combined)

	Federal Reserve	All Private forecasters
$(\hat{\alpha}_{0,j}^{h,k}, \hat{\delta}_{0,j}^{h,k})$ and $(\hat{\alpha}_{1,j}^{h,k}, \hat{\delta}_{1,j}^{h,k})$	$\frac{30}{84} = 0.357$	$\frac{131}{204} = 0.642$
$(\hat{\alpha}_{2,j}^{h,k}, \hat{\delta}_{2,j}^{h,k})$ and $(\hat{\alpha}_{3,j}^{h,k}, \hat{\delta}_{3,j}^{h,k})$	$\frac{37}{84} = 0.440$	$\frac{100}{204} = 0.490$

An interesting and significant difference between the Fed and private forecasters was the assessment of the effect of high interest rates on curbing inflation after 1979. Both the Fed as well as private forecasters underestimated the effect of Volcker's high interest rate regime on inflation in all horizons (i.e. $\hat{\alpha}_{2,2}^{h,k} < 0$, all h and all k). Private forecasters then turned around and overestimated the effect of monetary policy *in all horizons* during the Greenspan era (i.e. $\hat{\alpha}_{2,3}^{h,k} > 0$, all h and all k). That is, the private sector expected monetary policy under Greenspan to be more effective than it actually was. The Fed shared this mistake only in part. Its forecasters during the Greenspan era overestimated the effect of monetary policy (i.e. $\hat{\alpha}_{2,3}^{h,GB} > 0$) only for the short horizons of $h = 0$, 1.

4.4 *Testing the Accuracy of GDP Growth Forecasting (5b)*

We now present parameter estimates of the model (5b) for GDP growth forecast errors of the Fed, DRI, SPF and BLU. To some extent these are more important to real stabilization policy than inflation forecasting. The interpretation of the parameters is similar to the case of inflation forecasting and for this reason our presentation here is brief. Fed forecasts are available for 1965:11-1995:11, DRI's for 1970:7 -1995:11, SPF's for 1981:7-1995:11 and BLU's for 1980:1-1995:11.

4.4.1 *Assessing GDP Growth Forecast Accuracy*

We first recall the very low accuracy of GDP growth forecasting of the Fed and of private forecasters (see Table 3). We do not suggest forecasters should do better; we merely observe that accurate forecasting of GDP growth is very difficult. Hence, policy decisions of the Fed cannot be supported by precise forecasts of the impact of such decisions on future GDP growth. This implies that a central bank must recognize the market uncertainty cost of unexpected decisions. We also conclude that in the long run a central bank policy would attain a higher degree of stabilization if it is confined mostly to transparent and predictable policy rules.

We now turn to a brief examination of the pattern of forecast errors by the Fed forecasts of GDP growth reported in Table 7A. We make the following observations:

Table 7A: Estimated Fed's GDP Growth Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4	5	6
$\hat{\delta}^{h,GB}_{0,1}$	1.80	4.35*	4.18	2.74	-8.24	-16.4*	-21.7*
	(2.63)	(2.39)	(2.75)	(4.01)	(5.82)	(5.47)	(8.09)
$\hat{\delta}^{h,GB}_{0,2}$	.07	8.49*	8.29*	5.99*	4.82	.17	-.44
	(1.17)	(2.67)	(3.08)	(2.75)	(3.32)	(3.01)	(3.20)
$\hat{\delta}^{h,GB}_{0,3}$	1.57	3.52*	3.68*	5.04*	3.49	3.52	4.26
	(.99)	(1.43)	(1.67)	(1.82)	(2.43)	(2.84)	(3.22)
$\hat{\delta}^{h,GB}_{1,1}$	-.12	-.40*	-.43*	-.51	-.78	1.93*	2.58
	(.22)	(.18)	(.23)	(.40)	(.78)	(1.02)	(1.74)
$\hat{\delta}^{h,GB}_{1,2}$	.03	-.78*	-.90*	-.76	-.36	.50	.54
	(.11)	(.38)	(.51)	(.56)	(.73)	(.77)	(.71)
$\hat{\delta}^{h,GB}_{1,3}$	-.10	-.38	-.18	-.37	-.01	-.09	-.12
	(.13)	(.31)	(.37)	(.60)	(.74)	(.86)	(.88)
$\hat{\delta}^{h,GB}_{2,1}$	-.17	-.46*	-.44	-.22	.70	1.25*	1.76*
	(.32)	(.29)	(.34)	(.43)	(.52)	(.45)	(.51)
$\hat{\delta}^{h,GB}_{2,2}$	.10	-.57*	-.50*	-.31*	-.29*	-.08*	-.03*
	(.11)	(.19)	(.24)	(.17)	(.21)	(.18)	(.18)
$\hat{\delta}^{h,GB}_{2,3}$	-.11	-.36*	-.44*	-.57*	-.45*	-.52*	-.57*
	(.14)	(.20)	(.23)	(.25)	(.27)	(.25)	(.26)
$\hat{\delta}^{h,GB}_{3,1}$	-.16	-.11	-.08	-.36*	-.41*	-.15	.40
	(.21)	(.14)	(.18)	(.19)	(.18)	(.24)	(.47)
$\hat{\delta}^{h,GB}_{3,2}$	-.03	.02	- .05	.21	- .07	.11	- .06
	(.14)	(.17)	(.18)	(.23)	(.12)	(.27)	(.18)
$\hat{\delta}^{h,GB}_{3,3}$	.06	.13	.20*	.22	.24	.22*	.17
	(.09)	(.12)	(.13)	(.15)	(.16)	(.12)	(.13)
R^2	.04	.12	.08	.12	.10	.05	.07
N	293	277	255	238	208	149	89
Chi Square value for test: $\hat{\delta}^{h,GB}_{i,j}=0\ \forall i,j$	28.7	27.3	20.5	30.2	35.5	29.4	785.7

(i) As in the case of inflation, most point estimates satisfy $\hat{\delta}^{h,GB}_{0,j}>0$ and $\hat{\delta}^{h,GB}_{1,j}<0$: too low GDP growth forecasts when growth forecasts are slow and too high forecasts when forecasts are high.
(ii) In most cases $\hat{\delta}^{h,GB}_{2,j}<0$: Fed forecasters *underestimated* the effect of interest rates on future GDP

growth. *This strong result is consistent across different cases and very significant for policy.*

(iii) Known monetary policy variables explain a significant proportion of the Fed's forecast errors. The R^2 in Table 7A are around 10% but this fact is significant since the R^2 in Table 2 is very low: the variables in Table 7A improve significantly the Fed forecasting of GDP growth.

(iii) with 12 degrees of freedom the 10% critical value of a Chi Square test is 18.5 and the 1% value is 26.2 hence the joint perfect knowledge hypothesis $\hat{\delta}^{h,GB}_{i,j} = 0$ for all i and j is rejected for all h.

Table 7B: Estimated DRI's GDP Growth Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4	5	6	7
$\hat{\delta}^{h,DRI}_{0,1}$	3.77*	4.88*	5.91*	7.16*	6.42*	4.29	5.31	4.66
	(2.64)	(2.90)	(2.94)	(3.07)	(2.66)	(3.94)	(3.83)	(3.57)
$\hat{\delta}^{h,DRI}_{0,2}$	-.49	7.02*	7.26*	7.16*	7.18*	3.81*	4.75*	5.01*
	(1.19)	(1.89)	(2.33)	(2.03)	(1.86)	(1.52)	(2.12)	(2.05)
$\hat{\delta}^{h,DRI}_{0,3}$	2.66*	3.79*	5.04*	5.69*	4.84*	6.12*	6.93*	7.59*
	(1.19)	(1.59)	(1.65)	(2.03)	(1.88)	(1.73)	(1.70)	(1.60)
$\hat{\delta}^{h,DRI}_{1,1}$	-.23	-.34	-.45*	-.84*	-.95*	-1.02*	-1.86*	-2.28*
	(.20)	(.25)	(.23)	(.30)	(.17)	(.61)	(.65)	(.60)
$\hat{\delta}^{h,DRI}_{1,2}$	.12	-.45*	-.60*	-.76*	-1.04*	-.63*	-1.51*	-2 .22*
	(.10)	(.23)	(.37)	(.45)	(.35)	(.31)	(.46)	(.57)
$\hat{\delta}^{h,DRI}_{1,3}$	-.21	-.50	-.70*	-.74*	-.09	-.46	-1.12	-1.49*
	(.17)	(.26)	(.27)	(.41)	(.28)	(.39)	(.37)	(.38)
$\hat{\delta}^{h,DRI}_{2,1}$	-.39	-.62*	-.67*	-.62*	-.45	-.06	.41	.71*
	(.34)	(.31)	(.37)	(.36)	(.37)	(.34)	(.30)	(.29)
$\hat{\delta}^{h,DRI}_{2,2}$	.13	-.47*	-.48*	-.43*	-.33*	-.15	.06	.28*
	(.11)	(.17)	(.21)	(.22)	(.19)	(.16)	(.19)	(.17)
$\hat{\delta}^{h,DRI}_{2,3}$	-.20	-.31	-.45*	-.55*	-.64*	-.73*	-.60*	-.54*
	(.16)	(.21)	(.25)	(.27)	(.25)	(.24)	(.21)	(.19)
$\hat{\delta}^{h,DRI}_{3,1}$	-.02	.09	.13	-.31	-.57*	-.71*	-.72*	-.59*
	(.19)	(.16)	(.19)	(.20)	(.20)	(.28)	(.27)	(.23)
$\hat{\delta}^{h,DRI}_{3,2}$	.06	-.03	- .04	.15	- .13	-.11	-.19	-.43*
	(.12)	(.15)	(.16)	(.20)	(.17)	(.17)	(.17)	(.12)
$\hat{\delta}^{h,DRI}_{3,3}$	.12	.15	.19	.21	.36*	-.30*	.13	.13
	(.10)	(.12)	(.14)	(.14)	(.11)	(.12)	(.10)	(.11)
R^2	.05	.14	.15	.18	.25	.19	.24	.35
N	272	272	272	272	272	272	272	271
Chi Square value for test: $\hat{\delta}^{h,DRI}_{i,j} = 0 \; \forall i,j$	37.2	33.1	35.7	36.0	91.6	47.6	65.3	78.7

For DRI we make the following observations:

(i) In most cases $\hat{\delta}^{h,DRI}_{0,j} > 0$ and $\hat{\delta}^{h,DRI}_{1,j} < 0$, and here most parameters are statistically significantly different from zero. This is an important difference between DRI and the Fed.

(ii) $\hat{\delta}^{h,DRI}_{2,j} < 0$ for short horizons: DRI forecasters *underestimated* the effect of interest rates on future GDP growth. Some *overestimation* of the effect of policy for long horizons.

(iii) The R^2 are very large, increasing with the horizon. They are larger than in Table 7A for the Fed.

(iii) with 12 degrees of freedom the 10% critical value of a Chi Square test is 18.5 and the 1% value is 26.2 hence the joint perfect knowledge hypothesis $\hat{\delta}^{h,DRI}_{i,j} = 0$ for all i and j is rejected for all h.

For SPF and BLU in Tables 7C-7D we make the following joint observations:

(i) Most point estimates are $\hat{\delta}^{h,SPF}_{0,j} > 0$ and $\hat{\delta}^{h,BLU}_{0,j} > 0$ while $\hat{\delta}^{h,SPF}_{1,j} < 0$ and $\hat{\delta}^{h,BLU}_{1,j} < 0$. Similar to DRI, most are statistically significantly. This is a difference between private forecasters and the Fed.

(ii) In most cases $\hat{\delta}_{2,j}^{h,SPF} < 0$ and $\hat{\delta}_{2,j}^{h,BLU} < 0$: SPF and BLU forecasters *underestimated* the effect of interest rates on future GDP growth rate. Here again, monetary policy variables explain significant proportions of the SPF's and BLU's forecast errors with very high R^2.

(iii) with 8 degrees of freedom the 10% critical value of a Chi Square test is 13.5 and the 1% value is 20.1 hence each one of the joint perfect knowledge hypotheses $[\hat{\delta}_{i,j}^{h,SPF} = 0 \ , \forall \ i, \forall \ j]$ and $[\hat{\delta}_{i,j}^{h,BLU} = 0 \ \forall \ i, \forall \ j]$ are rejected for all h**.**

Table 7C: Estimated Median SPF GDP Growth Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4
$\hat{\delta}_{0,2}^{h,SPF}$	-2.02	-.17	3.36	6.03*	4.48
	(2.18)	(2.47)	(2.31)	(3.03)	(3.17)
$\hat{\delta}_{0,3}^{h,SPF}$	2.56*	4.30*	2.56*	2.56*	2.56*
	(1.62)	(2.26)	(1.62)	(1.62)	(1.62)
$\hat{\delta}_{1,2}^{h,SPF}$	.46*	1.39*	.93*	-1.04	-.72
	(.24)	(.43)	(.76)	(1.04)	(.63)
$\hat{\delta}_{1,3}^{h,SPF}$	-.09	-.32	-1.02*	-1.30*	-1.36*
	(.26)	(.43)	(.48)	(.43)	(.57)
$\hat{\delta}_{2,2}^{h,SPF}$	.18	-.44*	-.69*	-.19	-.13
	(.20)	(.14)	(.18)	(.23)	(.30)
$\hat{\delta}_{2,3}^{h,SPF}$	-.21	-.43*	-.02	-.70*	-.71*
	(.20)	(.27)	(.48)	(.31)	(.28)
$\hat{\delta}_{3,2}^{h,SPF}$	.12	-.56*	-.41*	-.27	-.23
	(.16)	(.30)	(.25)	(.32)	(.34)
$\hat{\delta}_{3,3}^{h,SPF}$	.13	.19	-.66*	.20	.24*
	(.13)	(.14)	(.31)	(.14)	(.15)
R^2	.01	.29	.27	.13	.11
N	78	77	76	75	74
Chi Square value for test: $\hat{\delta}_{i,j}^{h,SPF} = 0 \ \forall i, j$	23.8	33.6	29.4	27.8	37.2

Table 7D: Estimated Consensus BLU GDP Growth Forecast Error Functions

(standard errors in parentheses)

Horizon →	0	1	2	3	4	5	6
$\hat{\delta}_{0,2}^{h,BLU}$	1.14	6.68*	9.73*	9.45*	6.28*	5.65*	4.42*
	(1.60)	(2.60)	(2.72)	(2.59)	(1.94)	(2.28)	(2.75)
$\hat{\delta}_{0,3}^{h,BLU}$	2.88*	4.12*	8.05*	12.17*	10.39*	9.59*	10.64*
	(1.36)	(1.91)	(2.35)	(2.30)	(2.24)	(2.92)	(3.49)
$\hat{\delta}_{1,2}^{h,BLU}$	.06	-.40	-1.27*	-1.48*	-.83	-1.29*	-.43
	(.18)	(.38)	(.56)	(.76)	(.76)	(.68)	(.83)
$\hat{\delta}_{1,3}^{h,k}$	-.08	-.25	-1.23*	-2.48*	-2.05*	-1.66*	-2.25*
	(.27)	(.39)	(.50)	(.59)	(.62)	(.89)	(1.22)
$\hat{\delta}_{2,2}^{h,BLU}$	-.04	-.50*	-.51*	-.40*	-.35	-.10	-.23
	(.15)	(.19)	(.19)	(.21)	(.23)	(.23)	(.28)
$\hat{\delta}_{2,3}^{h,GB}$	-.27*	-.41*	-.69*	-.90*	-.77*	-.81*	-.81*
	(.16)	(.24)	(.29)	(.26)	(.26)	(.26)	(.27)
$\hat{\delta}_{3,2}^{h,BLU}$	.17	-.01	-.06	.14	-.32*	.00	-.05
	(.17)	(.17)	(.15)	(.18)	(.18)	(.22)	(.16)
$\hat{\delta}_{3,3}^{h,BLU}$	.15	.21*	.23*	.15	.15	.12	-.03
	(.11)	(.12)	(.13)	(.13)	(.13)	(.13)	(.17)
R^2	.02	.13	.25	.27	.28	.15	.18
N	252	249	246	243	235	174	114
Chi Square value for test: $\hat{\delta}_{i,j}^{h,BLU} = 0 \ \forall i, j$	29.4	28.1	43.9	88.4	65.9	32.7	27.3

4.4.2 *Some Additional Notes on the Accuracy of GDP Growth Forecasts*

Our results for GDP growth complement the results for inflation forecasting. These are:

(i) All GDP growth forecasts are very poor and Although monetary policy variables improve GDP growth forecasts, practically speaking the improvement is not sufficient.

(ii) Composite Wald tests of the perfect knowledge hypothesis are universally rejected by the data.

(iii) The common element of the error functions reviewed in Tables 7A-7D consist of two parts:

- in most cases $\hat{\delta}_{0,j}^{h,k} > 0$ and $\hat{\delta}_{1,j}^{h,k} < 0$: forecasts of GDP growth were too low when growth forecasts were low and GDP growth forecasts were too high when forecasts were high.
- in almost all cases $\hat{\delta}_{2,j}^{h,k} < 0$: forecasters almost universally underestimated the stabilization effect of interest rates on GDP growth. In some cases we find overestimation for long horizons.

(iv) A key difference between the Fed and private forecasts is that ($\hat{\delta}_{0,j}^{h,GB}$, $\hat{\delta}_{1,j}^{h,GB}$) are not significantly different from zero for the Fed but for private forecasters $\hat{\delta}_{0,j}^{h,k}$ and $\hat{\delta}_{1,j}^{h,k}$ are different from zero. This implies that, by and large, Fed forecasters erred mostly in predicting the effect of monetary policy on GDP growth while errors associated with other variables are generally small. Private sector forecasters erred in assessing the impact of both monetary policy variables as well as other economic variables.

Conclusion (iii) is of particular interest. Why is it that, for almost all horizons, all forecasters *underestimated* the stabilization effects of interest rates on GDP growth? Contrast this finding with the fact that a general underestimation of the effect of monetary policy was true of inflation forecasting *only during the Volcker monetary regime*. In the first and third regimes forecasters *overestimated* the effect of monetary policy on inflation, particularly at short horizons. In order to clarify what the data tells us, we recall the evidence of correlation in the subjective models used by forecasters. This means that similar concepts, ideas and assumptions underlie the models which diverse forecasters use and *this fact shows best in the pattern of their mistakes*. To understand the pattern we thus need to examine the ideas about monetary policy which forecasters may have held. One may conjecture that our finding reflects the impact of the monetarist debate in the Economics profession over the last thirty years. If academic writings insist that monetary policy cannot have any systematic real effects on GDP growth, how would this affect the Fed and private forecasters? They would tend to build into their subjective models assumptions that would underestimate the effect of money on the real economy. With ideas of money neutrality and the irrelevancy of monetary policy dying slowly, forecasters tended to learn very slowly from their past forecasting errors.

A corresponding argument explains the pattern of inflation forecast errors. The Volcker monetary regime was initiated as an aggressive anti inflationary drive during time when inflation expectations took firm hold in the U.S. A skeptical view of the Volcker policy resulted in universal underestimation of the ability of high interest rate to suppress inflation. But once the Volcker policy did succeed in lowering inflation, private forecasters became so convinced of the ability of the Fed to control inflation that they *overestimated* the inflation stabilization ability of the Greenspan Fed in all forecast horizons. This overestimation was shared by the Fed, but only for short horizons.

References

Barrow, R. J., Gordon, D. B. (1983): Rules, Discretion and Reputation in a Model of Monetary Policy." *Journal of Monetary Economics*, **12** , 101-121.

Cochrane, J., Piazzesi, M. (2002): "The Fed and Interest Rates - A High-Frequency Identification." *American Economic Review* **92**, 90 - 95.

Ehrbeck, T., Waldman, R. (1996): "Why are Professional Forecasters Biased? Agency versus Behavioral Explanations." *Quarterly Journal of Economics* **111**, pp. 21-40.

Ellingsen, T., Söderström, U. (2001): "Classifying Monetary Policy." Sveriges Riksbank Working Paper 56.

Faust, J., Swanson, E., Wright, J. (2004): "Do Federal Reserve Policy Surprises Reveal Superior Information About the Economy?" *Contributions to Macroeconomics*, The Berkeley Electronic Press, **4,** pp. 1- 29.

Kurz, M. (1974): The Kesten-Stigum Model and the Treatment of Uncertainty in Equilibrium Theory. In: Balch, M.S., McFadden, D.L., Wu, S.Y., (ed.), *Essays on Economic Behavior Under Uncertainty*, pp. 389-399, Amsterdam: North Holland.

Kurz, M. (1994): On the Structure and Diversity of Rational Beliefs. *Economic Theory* **4**, pp. 877 - 900 . (An edited version appears as Chapter **2** of Kurz, M. (ed.) (1997))

Kurz, M. (ed) (1997): *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*. Studies in Economic Theory, No. **6**, Berlin and New York: Springer-Verlag.

Kurz, M. (2002): "Heterogenous Forecasting and Federal Reserve Information." Working paper No.02-002, Department of Economics, Stanford University.

Kurz, M., Jin, H., Motolese, M. (2005a): "Determinants of Market Volatility and Risk Premia." Working paper No. 04-001, Department of Economics, Stanford University. To appear in *Annals of Finance* **1** .

Kurz, M., Jin, H., Motolese, M. (2005b): "The Role of Expectations in Economic Fluctuations and the Efficacy of Monetary Policy." Working paper No. 03-010, Department of Economics, Stanford University. To appear in the *Journal of Economic Dynamics and Control*.

Kydland, F. E., Prescott, E. C. (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, **85**, 473 - 491.

Newey, W.K., West, K.D. (1987): "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55, pp.703-708.

Romer, C. D., Romer, D. H. (2000): "Federal Reserve Information and the Behavior of Interest Rates." *American Economic Review*, June, pp. 429-457.

Scharfstein, D. S., Stein, J. C. (1990): "Herd Behavior and Investment," *American Economic Review*, **80**, pp. 465 - 479.

Shaffer, S. (1998): "Information Content of Forecast Errors," *Economic Letters*, **59**, pp. 45 - 48.

Shaffer, S. (2003): "Using Prior Bias to Improve Forecast Accuracy," *Applied Economic Letters*, **10**, pp. 459 - 461.

Sims, C. (2002): "The Role of Models and Probabilities in the Monetary Policy Process." *Brookings Papers on Economic Activity* **2**, 1- 62.

Stock, H. J., Watson, W. M. (1999): "Forecasting Inflation." Department of Economics, Princeton University.

Stock, H. J., Watson, W. M. (2003): "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, **41**, 788 - 829.

Stock, H. J., Watson, W. M. (2002): "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics*, The American Statistical Association, **20**, 147-162.