Appendix to: Fiscal Forecasts at the FOMC: Evidence from the Greenbooks

By Dean Croushore and Simon van Norden

This appendix describes some research that we have left out of the main paper. It includes more details on data definitions, more detailed results of the Romer-Romer regressions, the results of testing a fiscal-policy influenced Taylor rule, and the results of Patton-Timmermann tests.

I. Data Definitions

A. Sources and Sample

The Greenbook is summary of economic conditions, trends and forecasts prepared for every meeting of the FOMC. Our primary data sources are page scans of each Greenbook made available by the Board of Governors of the Federal Reserve System[1] and by the Real Time Data Research Center at the Federal Reserve Bank of Philadephia.[2] These two sources provide independently-made page scans from different physical copies of the vintage historical materials; this allowed us to independently confirm figures which, on a few very rare occasions, were difficult to distinguish or missing in one of the two sources.[3]

The first data vintage collected was for July 1966 and the last was December 2006, covering 387 meetings of the FOMC over 40 years. This represented the full set of source materials available when we started. However, the earliest versions either lack fiscal variables or contain only very short time series (typically five quarters, most of which are historical estimates.) Most of our fiscal variables (Surplus, Revenues and Expenditures) first appeared in the August 1967 Greenbook while the first appearance of the HEB variable was in April 1970. FOMC meeting dates are slightly irregular, but for most of the period there were exactly two meetings per quarter. Meetings in the early part of the sample were more frequent (12 or more per year, but not necessarily one per month.) The release dates of key statistics also vary somewhat over the years. To standardize the forecast horizons we examine, we restrict our analysis to the vintages from the first and the last FOMC meeting of each quarter. A complete list of data vintage

[1] `www.federalreserve.gov/monetarypolicy/fomc/`

[2] `www.philadelphiafed.org/research-and-data/real-time-center/greenbook-data/`

[3] Note that the Greenbook estimates published in the ALFRED database at the Federal Reserve Bank of St. Louis only contains figures from the main volumes of the Greenbook. This is compiled a few days prior to the meeting of FOMC; late-breaking developments (such as statistical releases or revisions) are collected and circulated in the form of a supplement to the Greenbook. Our data reflect the estimates presented to the FOMC; these incorporate any additions or revisions contained in supplements to the Greenbook.

dates is provided below and are summarized in the following Table.

Table 1—Available Data Vintages (FOMC Meeting Dates)

Series	First Vintage	Last Vintage
Surplus	August 1967	December 2006
Receipts	August 1967	December 2006
Expenditures	August 1967	December 2006
Unemployment	July 1966	December 2006
GDP (nominal)	July 1966	December 2006
GDP (real)	July 1966	December 2006
HEB	April 1970	December 2006
HEB (6%)	November 1980	December 2006

The number of observations and the forecast horizons included in each series varied considerably over time. Our figures were principally compiled from the *Federal Sector Accounts* and *Main Economic Indicators* Tables (whose contents varied somewhat over the years.) When series were shown in both tables, we collected data from both to maximize the span of observations available. In some of the earliest vintages, series might not contain more than 5 Q of historical estimates and forecasts, whereas later vintages could contain up to 20Q. Greenbooks often had slightly more quarters of historical estimates than of forecasts, as can be seen in Table 2, which gives one example of the number of available forecasts for each forecast horizon.

B. Validation

The data were validated in a number of ways.

1.) A professional data-entry firm was employed for initial key-input of the data with a contracted accuracy rate $\geq 99.95\%$.

2.) Several of their series were then checked against independent sources. This verified the claimed accuracy rate.

2a) Unemployment rates, as well as nominal and real levels of GNP and GDP were checked against estimates published in ALFRED by the FRB St. Louis. We found 10 cases where the figures in ALFRED did not correspond to the page scans, and one case where we had missed an entry.[4] We also found a number of cases where the FOMC and the FRB Philadelphia page scans disagreed. In those cases, the FRB Philadelphia page scans were dated slightly after the original Greenbook estimates, indicating that figures were revised just prior to the FOMC meeting. We used the latter.

[4]We communicated our findings to the FRB St. Louis, who verified our figures and corrected the entries in ALFRED. Note that with slightly more than 5000 data points checked, this implies a pre-correction error rate for Greenbook series in ALFRED of $< 0.2\%$ and $< 0.02\%$ for our data entry.

Table 2—Number of Observations by Forecast Horizon: Govt. Receipts (Outcome = First Reported Value)

Forecast Horizon	First Meeting	Last Meeting	Forecast Horizon	First Meeting	Last Meeting
-12	0	0	**12**	0	0
-11	2	0	**11**	0	0
-10	3	2	**10**	0	0
-9	3	3	**9**	1	6
-8	3	3	**8**	9	17
-7	3	3	**7**	18	26
-6	16	3	**6**	39	52
-5	40	19	**5**	65	74
-4	73	49	**4**	88	101
-3	120	92	**3**	117	128
-2	152	143	**2**	135	137
-1	158	157	**1**	146	150
0	158	158			

2b) HEB estimates were checked against estimates entered independently.[5] Of approximately 3,000 data points, we found and corrected 10 discrepencies (0.3%); three were due to incorrect or missing meeting dates, five were due to keying errors in the independent estimates, and the remainder due to illegible page scans.

3.) There were a small number of cases in which figures shown in the *Federal Sector Accounts* Table were not precisely the same as those shown in *Main Economic Indicators* Table of the same Greenbook. One possibility is that the two tables may have been prepared by different groups; older Greenbooks were compiled by hand and slight discrepancies may have arisen in preparation.

4.) We verified that the Surplus/Deficit data were consistent with the data for Receipts and Expenditures.[6]

C. Forecasts

We recorded all Greenbook estimates for our selected series. This included estimates for future periods (forecasts), current periods (nowcasts) and historical periods (backcasts.) In this appendix, we collectively refer to all of these as *forecasts* although some prefer the term "projection" to emphasize the conditional nature of these estimates. Forecast horizons varied widely from meeting to meeting. At times, the convention was that the forecast horizon was fixed to the end of a given calendar year, then rolled forward once a year. This meant that

[5]The authors would like to thank Wendy Chan of the Bank of Canada for her research assistance.

[6]Figures in the Greenbook for May 1999 incorrectly reversed the sign on the Deficit. We corrected the sign.

the length of the forecast horizon varied somewhat through the year. There was also a general tendency for forecast and backcast horizons to increase across the decades, although there were some occasions when the horizons were decreased (perhaps because the longest horizons were not felt to be useful.) When series were listed in more than one table, different tables might include different forecast horizons. As the content of the tables evolved over time, the available forecast horizons might therefore vary from series to series.

D. Outcomes

Forecast evaluation requires a measure of observed outcomes. One of the series we collect (HEB) has no officially published value; it is only calculated by Board staff. While the other series correspond to official statistics, values published for the latter are revised over time. These revisions may reflect the incorporation of new information as preliminary published estimates are refined in the quarters immediately following their initial publication. It may also reflect conceptual changes in the definition of the series, such as the change from GNP to GDP or from a fiscal surplus to a fiscal current account surplus. We refer to the latter as "benchmark" revisions. Each of our series were affected, to greater or lesser degress, by benchmark changes. This complicates the measurement of forecast outcomes. We therefore use a variety of different "outcome" concepts to provide alternative characterizations of forecast performance. They are

First Release: This is the initial quarterly estimate published by the responsible official statistical agency (BEA or BLS.)

One Year: This is the official quarterly estimate that was available precisely one year after the publication of the First Release. For example, if the First Release was published on 23 September 1998 and revisions were published on 26 August 1999 and 29 September 1999, the August 1999 estimate would be the One Year estimate. This typically incorporates the annual revision common to most official series.

Last Greenbook: This is the last value recorded in the Greenbook, typically one or more years after the quarter to which it refers. This is primarily important as a measure for HEB, which has no counterpart in official statistics.

Pre-Benchmark: This is the last official estimate reported prior to a benchmark revision of the series. This is intended to capture the most precise available estimate of the same concept that the staff were forecasting and has previously been used in the literature as a measure of data revision.[7]. We discuss the identification and importance of benchmark revisions below.

Final: This is a "contemporary" estimate, which in our case was the official estimate as of December 27, 2012.

[7]For example, see Aruoba (2008).

E. Benchmark Revisions

We use the extent of revision to define those which we treat as *benchmark* revisions. We treat as benchmark revisions those which affect the entire published history of a time series. For example, US Quarterly National Accounts are available starting from 1946Q1. Revisions which do not affect the published estimates for the 1940s are therefore not considered benchmark revisions. Changes in seasonal adjustment factors, although they may occur many years after the fact, are not counted as benchmark revisions. Changes in base years (for real values), or the change from fixed-weight to chain-weighted values, or the change from GNP to GDP, are all examples of benchmark changes. This definition of benchmark revision has at least two important advantages.

1.) It is a simple, transparent and objective way to determine which revisions are be treated as benchmark revisions.

2.) It implicitly relies on the judgement of the statistical agency to determine which methodological or conceptual changes are important enough to be considered benchmark changes. In effect, if the statistical agency judges that historical estimates are sufficiently comparable to current estimates that no revision to the former is required, no benchmark revision has occured.

This definition also has at least one important drawback: since no official series is published for HEB, no long time series are available to identify benchmark changes. As we describe below, we therefore treat HEB estimates somewhat differently.

The economic importance of benchmark revisions varied vastly across our series, as we describe below in greater detail. At one extreme, benchmark revisions in the unemployment rate were rare and trivial. In contrast, the redefinition of the government accounts had an important impact on our fiscal variables. We discuss the economic importance of benchmark revisions in the next subsection. Table 3 shows the dates at which benchmark revisions were first published for each series.

Values forecast prior to benchmark revision are not comparable to outcomes measured after a benchmark revision. For that reason, whenever a forecast or nowcast is made for an outcome that will only be observed after a benchmark revision has occurred, we drop those forecast errors from our data set. For example, the Greenbook for the FOMC meeting in October 1975 contained nowcasts and forecasts for the period 1975Q4-1976Q4. Estimates for most of these outcomes were only published after the benchmark revision which was first released on January 20, 1976. Therefore, for the series affected by those benchmark changes, those forecast errors were replaced by a missing value code.

F. Variables

GNP & GDP: Our outcome measures for these series were taken from ALFRED series *GNP* and *GDP.* The BEA published estimates of GNP until December 1991, after which it switched to GDP as its main measure of economic

Table 3—Pre-Benchmark-Revision Dates for Quarterly National Accounts

Last Quarter	Last ALFRED Vintage	Last FOMC Date
1975:3	Dec. 19, 1975	Dec. 10, 1975
1980:3	Nov. 19, 1980	Dec, 12, 1980
1985:3	Nov. 20, 1985	Dec. 11, 1985
1991:2	Aug. 28, 1991	Oct. 30, 1991
1995:2	Oct. 27, 1995	Dec. 14, 1995
1999:2	Sep. 30, 1999	Sep. 29, 1999
2003:3	Nov. 25, 2003	Dec. 03, 2003

Note: This table gives the dates of publication for the last estimates prior to benchmark revisions of the National Accounts. The first column gives the last time period to which those estimates correspond. The second column gives the date at which those estimates were published. The last column gives the date of the last FOMC meeting prior to the publication of the benchmark revision. These dates apply to figures from the Quarterly National Accounts as based on original data vintages from ALFRED and the FRB Philadelphia Real-Time Data Set for Macroeconomists. The 1995 benchmark revision of Expenditures occurred slightly after the revision of the other series; its last pre-benchmark-revision quarter was 1995:3 which was published on October 27, 1995. The last FOMC meeting using this estimate was that of December 1995.

activity. The Greenbooks followed suit, focusing on GNP until that date and GDP thereafter. Our primary use of these series is to express various fiscal series as a fraction of the overall size of the US economy, for which we need an estimate of the level of the series. After August 2005, Greenbooks no longer list GDP in levels, giving only growth rate forecasts. For the last 11 FOMC meetings we recorded, we therefore calculated an implied level GDP forecast from the growth rate forecasts by applying the compound growth rate to the second-to-last (and therefore already revised) officially published estimate. For example, the growth rate estimates from the September 2005 Greenbook are applied to the August 31, 2005 vintage BEA estimate of GDP. The last estimate in that vintage is for 2005Q2; we therefore use the 2005Q1 estimate of 12198.8 as our base.

Receipts, Expenditures and Surplus/Deficit: Outcomes for the Surplus/Deficit were measured by ALFRED series *FGDEF: Net Federal Government Saving*. Outcomes for Receipts were taken from *FGRECPT: Federal Government Current Receipts*, and for Expenditures from *FGEXPND: Federal Government: Current Expenditures.*

HEB: The *High-Employment Budget Surplus/Deficit* (HEB) is the Greenbook's estimate of a cyclically-adjusted or "structural" budget deficit. This is the Board staff's counterfactual estimate of what the surplus (or deficit) would be if the unemployment rate were at a constant reference level over the forecast horizon. The budget deficit concept used in HEB always corresponds to that used in the Surplus/Deficit measure; prior to 1996 this was the overall Surplus or Deficit, and this was replaced by the Government Current and Capital Account Surplus/Deficit thereafter.

The reference level of unemployment used to calculate HEB is not always explicitly mentioned, but drifted upwards from near 4.0% in the earliest part of

our sample before major changes were introduced in 1980. From November 1980 until March 1983, two alternative HEB estimates were presented, based on a 6.1% and a 5.1% reference level of unemployment. From May 1983 until August 1983 these were replaced by rates of 6.0% and 5.0%. Thereafter, the reference level was constant at 6.0%. We assume that these changes reflected uncertainty and disagreement within the Board about the natural rate of unemployment. The Table design during the "dual-rate" period gave greater prominence to the 6.1% (and then the 6.0%) reference level.

We found that the revision of the reference level of unemployment appeared to have a qualitatively important effect on the HEB estimates. We therefore consider two different sets of HEB estimates; the full series as well as the subset (HEB6) which only considers those estimates based on a 6.0% or 6.1% reference level. We make no attempt to adjust the HEB6 series for the change from 6.1% to 6.0%. We also calculate the difference between the HEB (and HEB6) estimates and the overall Surplus/Deficit estimates as the Board Staff's implied estimate of the cyclical Surplus/Deficit.

Unemployment: Outcomes for this series were measured by ALFRED series *UNRATE: the Civilian Unemployment Rate.* Greenbooks only report the unemployment rate to one decimal place. Starting with the official estimate published on Feb. 9, 1967, the labor force was redefined to count only those age 16 and over instead of 14 and over. This never caused revisions of more than 0.1% in absolute value in our data set. There were no benchmark revisions to unemployment after that date. We therefore chose to ignore benchmark revisions in the unemployment rate and do not use a "Pre-Benchmark" measure of outcomes.

II. Romer-Romer Regression Detailed Results

Table 4 compares estimates of the Romer and Romer (2004) original regression with those incorporating Greenbook estimates of the Federal Government Surplus. The first pair of columns simply repeats the Romer's original work, regressing the changes in the Federal Funds Rate Target on their selected control variables.[8] In the following pair of columns, we add Greenbook forecasts for current and future values of the Surplus/Deficit.[9] In addition to an economically important rise in the R^2, the added variables are jointly statistically significant while some of the variables in the Romers' original specification no longer appear to be. We are also unable to reject the null hypothesis that the sum of the coefficients on the Surplus equals zero.

[8]We use the Romers' original data set, which the authors have graciously made publicly available. Throughout our analysis, our regressions are estimated by Ordinary Least Squares. Although our data set includes the all the FOMC meetings used in the Romers' study, many of the earliest meetings contain forecasts for only the very shortest horizons. Including a four-quarter forecast horizon for our Surplus variable reduces the sample from the Romers' original 261 observations to 220; including a four-quarter forecast horizon for HEB6 further reduces this to 122.

[9]Revisions of the fiscal forecast were also examined, but were never statistically significant.

We therefore replace the Surplus with its own first difference and drop sets of insignificant variables, which leads us to our preferred specification in the next pair of columns. The result is an equation with fewer explanitory variables and a slightly higher R^2 than the original regression, as well as a larger (more negative) coefficient on the old Target rate. As the latter controls the rate at which policy shocks decay, this implies that monetary policy shocks have somewhat less persistent effects on the target rate. We also note that both specifications show that expected fiscal policy 3-4 quarters into the future has a statistically significant effect on the Federal Funds Rate, with larger Surpluses associated with lower interest rates as one might expect from a monetary policy aiming to stablize aggregate demand.[10]

The following pair of columns uses HEB6 in place of the Surplus. While the structural surplus may be a more conceptually appealing as an indicator of the fiscal policy stance, it constrains our sample period to begin only in November 1980, thereby reducing our availabe number of observation by roughly half. However, we again find that this fiscal variable appears to play a statistically significant role in our shortened sample while one of the Romer and Romer variables (the change in forecast inflation) does not. Dropping the latter leaves us with our preferred specification for the structural deficit. We note that, in addition to coefficients signs similar to those we discussed previously, the overall fit of the equation over this sample is much better than that over the full sample, suggesting that exogenous monetary policy shocks were relatively less important. This may in part reflect the omission of large changes in the rate in 1979 and 1980, including 387.5 basis point movement on April 22, 1980, which is by far the largest movement in the sample. We also see a much larger (more negative) coefficient on the lagged level of the target, suggesting that monetary policy shocks were much less persistent.

III. A Fiscal Policy-Influenced Taylor Rule

Economists have often discussed the interactions between monetary policy and fiscal policy, yet simple rules such as the Taylor rule described in Taylor (1993) do not include a variable representing fiscal policy. Given that the Fed's Greenbook provides a substantial amount of information on fiscal policy, evidently monetary policymakers consider fiscal policy details while determining their policy actions.

So, suppose we estimate a Taylor rule and include a term representing fiscal policy. Would such a term be significant, and would it have an impact on monetary policy? We can use the Greenbook forecasts to form a Taylor rule that either excludes or includes forecasts of the overall and the cyclical part of the government budget surplus.

The Taylor Rules we estimate take the general form

[10]However, we also note the counterintuitive result that fiscal policy at other forecast horizons is also statistically significant but with the opposite sign.

TABLE 4—REVISED ESTIMATES OF ROMER AND ROMER (2004)

	Coefficient	t-Ratio	Coefficient	t-Ratio	Coefficient	t-Ratio	Coefficient	t-Ratio	Coefficient	t-Ratio
Constant	0.169	1.177	0.128	0.600	0.068	0.878	- 0.022	- 0.088	- 0.041	- 0.208
OLDTARG	- 0.021	- 1.751	- 0.029	- 1.745	- 0.027	- 2.518	- 0.066	- 2.759	- 0.058	- 2.515
GRAYM	0.007	0.667	- 0.005	- 0.380			0.022	1.300	0.024	1.502
GRAY0	0.003	0.174	0.008	0.324			0.009	0.297	- 0.007	- 0.239
GRAY1	0.010	0.304	0.006	0.159			0.092	1.794	0.084	1.670
GRAY2	0.022	0.685	0.039	0.893			- 0.059	- 0.841	- 0.065	- 1.222
IGRYM	0.049	1.611	0.033	1.007	0.019	0.678	- 0.016	- 0.414	- 0.028	- 0.748
IGRY0	0.152	5.030	0.140	4.102	0.143	5.444	0.093	2.349	0.101	2.636
IGRY1	0.021	0.454	0.022	0.393	0.046	1.120	- 0.033	- 0.478	- 0.021	- 0.311
IGRY2	0.017	0.331	0.017	0.263	0.031	0.679	0.186	2.026	0.227	2.941
GRADM	0.021	0.886	0.009	0.289	0.024	1.255	- 0.022	- 0.571	- 0.001	- 0.016
GRAD0	- 0.044	- 1.490	- 0.091	- 2.482	- 0.055	- 2.053	- 0.038	- 0.795	- 0.045	- 1.130
GRAD1	0.010	0.229	0.005	0.094	0.017	0.382	0.106	1.430	0.060	1.018
GRAD2	0.051	1.072	0.115	1.793	0.040	0.861	0.125	1.516	0.140	2.002
IGRDM	0.058	1.284	0.057	1.149			0.029	0.547		
IGRD0	0.002	0.036	0.033	0.582			- 0.031	- 0.497		
IGRD1	0.028	0.378	- 0.005	- 0.053			- 0.161	- 1.550		
IGRD2	- 0.065	- 0.792	- 0.103	- 1.027			- 0.072	- 0.579		
GRAU0	- 0.047	- 2.242	- 0.017	- 0.333			- 0.011	- 0.298		
SRPL0			- 8.614	- 1.442						
SRPL1			24.972	2.669						
SRPL2			15.086	1.217						
SRPL3			- 38.615	- 2.366						
SRPL4			9.921	0.765						
DSRPL1					26.010	3.370				
DSRPL2					5.916	0.608				
DSRPL3					- 20.214	- 1.968				
HEB0							4.575	0.527	6.617	0.773
HEB1							- 6.555	- 0.471	- 11.815	- 0.881
HEB2							36.017	1.522	43.767	1.911
HEB3							- 80.482	- 3.278	- 86.200	- 3.830
HEB4							54.974	4.018	54.915	4.851
R^2	0.280		0.351		0.289		0.503		0.481	
# Obs	261		220		237		122		103	
$[FISCAL] = 0$			3.780		5.327		4.442		5.646	
under H_0			F(5,196)		F(3,224)		F(5,98)		F(5,103)	
p-Value			0.003		0.001		0.001		0.000	

Notes:

OLS regression results for dependent variable *DTARG* - the change in the Federal Funds Rate Target.
OLDTARG - Federal Funds Rate Target before start of FOMC meeting.
GRAYM-GRAY2 - Greenbook forecast rate of output growth (M=previous quarter, 2=2 quarters from now).
IGRYM-IGRY2 - Change in Greenbook forecast rate of output growth from previous FOMC meeting (M=previous quarter, 2=2 quarters from now.)
GRADM-GRAD2 - Greenbook Inflation forecast (M=previous quarter, 2=2 quarters from now.)
IGRDM-IGRD2 - Change in Greenbook inflation forecast from previous FOMC meeting (M=previous quarter, 2=2 quarters from now.)
GRAU0 - Greenbook Unemployment Rate estimate for current quarter
SRPL0-SRPL4 - Greenbook forecast Surplus(Deficit) to GDP ratio (0=current quarter, 4=4 quarters from now.)
DSRPL1-DSRPL3 - Greenbook forecast change in Surplus(Deficit) to GDP ratio (1=next quarter, 3=3 quarters from now.)
HEB0-HEB4 - Greenbook forecast High-Employment Budget Surplus (Deficit) to GDP ratio (0 = current quarter, 4=4 quarters from now.)

[FISCAL]=0 - F-test of the null hypothesis that all coeficients on the fiscal variables are zero.

$$i_t = \rho i_{t-1} + (1-\rho)\hat{i}_t + \epsilon_t.$$

$$\hat{i}_t = \beta' X_t,$$

where X_t is a vector of variables known at time t, or forecasts formed at date t of variables at future dates. Early versions of the Taylor rule had lagged values of the inflation rate and the output gap in the vector X but we will consider a wider range of variables, including forecasts of the future inflation rate, the output gap, and fiscal variables.[11]

What variables should be included in vector X? The standard Taylor rule includes just the output gap and the difference between the inflation rate and the target for inflation. Empirical estimation assumes that the inflation target is constant over time, so the independent variables in a regression analysis include a constant term, the output gap, and the inflation rate. The other major modeling choice is whether the terms should be observables (lagged output gap and inflation rate) or forecasts (forecasts of output gap and inflation rate), reflecting the forward-looking nature of monetary policy. We show results for the Taylor rule based on observables (data at time $t-1$), current-quarter forecasts, two-period-ahead forecasts, and four-period-ahead forecasts.[12]

If the Fed pays attention to fiscal variables in setting monetary policy, how would those variables be reflected in the Taylor rule? One possibility is that the fiscal variables are useful only for determining the output gap, in which case they should not enter the Taylor rule separately. But it may be that fiscal variables influence monetary policy directly rather than just indirectly via the output gap. In that case, we would expect the addition of fiscal variables to significantly improve the fit of the Taylor Rule. To examine this issue, we will estimate the Taylor rule including terms for both the overall surplus (S_{t+k}) and structural surplus ($HEB6_{t+k}$), separately. The results are shown in Table 5.

All of the estimated coefficients on inflation, the output gap and lagged interest rates have the expected sign. The coefficient on i_{t-1} is always strongly significant with coefficient estimates in the range of 0.73 to 0.87, which implies that the Fed reacts to shocks fairly slowly. The coefficients on π_{t+k} are always positive and

[11] Rudebusch (2006) finds some evidence favoring rules of the form

$$i_t = \beta' X_t + \nu_t,$$

$$\nu_t = \gamma \nu_{t-1} + \omega_t.$$

where unmodeled shocks ν_t are serially correlated. We estimated several such models and found that they tended to give somewhat similar, albeit weaker, results. While Rudebusch (2006) argues that the two forms are difficult to distinguish empirically, Coibion and Gorodnichenko (2012) present additional evidence that strongly favors the first form presented above. For that reason, we present only those results.

[12] We estimated the rule using 1-quarter-ahead and 3-quarter-ahead forecasts, but do not report these results to conserve space. Those results contribute no additional insights.

TABLE 5—TAYLOR RULE ESTIMATES

	k = -1				k = 0			
	First		Last		First		Last	
Constant	0.247	0.599	0.249	0.440	0.298	0.690	0.260	0.633
	(1.28)	(2.33)	(1.44)	(1.82)	(1.81)	(3.70)	(1.56)	(3.25)
i_{t-1}	0.865	0.836	0.856	0.836	0.815	0.820	0.857	0.864
	(25.3)	(24.5)	(26.6)	(27.1)	(27.3)	(25.7)	(33.2)	(28.1)
π_{t+k}	0.179	0.091	0.195	0.115	0.263	0.136	0.181	0.053
	(3.68)	(1.69)	(4.33)	(1.92)	(3.11)	(1.82)	(2.70)	(0.86)
y_{t+k}	0.138	0.377	0.153	0.359	0.179	0.393	0.164	0.384
	(4.69)	(6.15)	(5.90)	(5.62)	(7.39)	(8.56)	(7.50)	(8.26)
S_{t+k}		-38.2		-31.2		-39.7		-40.7
		(3.40)		(2.44)		(5.72)		(5.23)
$HEB6_{t+k}$		26.9		18.1		32.6		32.3
		(2.57)		(1.47)		(4.62)		(3.94)
$\overline{R}^2$	0.961	0.973	0.964	0.975	0.968	0.980	0.965	0.978
	k = 2				k = 4			
	First		Last		First		Last	
Constant	0.208	0.551	0.181	0.466	0.038	0.513	0.009	0.438
	(1.48)	(2.67)	(1.31)	(2.42)	(0.25)	(2.05)	(0.06)	(1.89)
i_{t-1}	0.733	0.793	0.771	0.821	0.781	0.822	0.809	0.841
	(26.3)	(28.3)	(30.3)	(38.9)	(22.0)	(29.6)	(25.2)	(27.3)
π_{t+k}	0.484	0.268	0.412	0.240	0.458	0.267	0.406	0.261
	(6.50)	(3.04)	(5.82)	(3.88)	(4.34)	(2.44)	(4.49)	(2.57)
y_{t+k}	0.255	0.382	0.223	0.340	0.231	0.402	0.206	0.344
	(8.00)	(7.67)	(7.99)	(7.42)	(5.61)	(5.97)	(6.39)	(5.74)
S_{t+k}		-30.1		-27.4		-34.4		-29.2
		(3.32)		(3.14)		(3.21)		(2.68)
$HEB6_{t+k}$		26.7		24.8		34.3		32.0
		(2.74)		(2.49)		(3.03)		(2.69)
$\overline{R}^2$	0.977	0.981	0.976	0.979	0.973	0.978	0.972	0.975

Note: S_{t+k} refers to the Surplus and $HEB6_{t+k}$ refers to the structural surplus. Coefficients are shown for each variable in the Taylor Rule equation, with t-statistics shown in parentheses below each coefficient. Estimation is by least squares with HAC standard errors. *First* and *Last* refer to the timing of the FOMC meeting within the quarter.

almost always significant, as we expect. They increase in size as k increases, as we would expect if the Fed reacts more strongly to expected future inflation than to past inflation. Coefficients on the output gap (defined as the level of output minus the level of potential output,expressed as a percentage of the latter) are always of the correct sign and strongly significant.

The two fiscal balance variables are almost always both statistically significant at the 5 percent confidence (the only exception being the case except when $k = 1$ in the last meeting of the quarter for $HEB6$.) The sign on the overall budget surplus is negative, implying that the Fed eases monetary policy when fiscal policy tightens. But the coefficient on $HEB6$ is positive, which implies that a higher structural surplus leads the Fed to *tighten* monetary policy. The introduction of these variables also always raises the coefficient on the output gap and lowers that on lagged inflation, sometimes to the point of insignificance.

If the economy is projected to be at full employment, then $S_{t+k} = HEB6_{t+k}$, so the direction of the fed funds rate would be determined by the sum of the coefficients on S_{t+k} and $HEB6_{t+k}$. In cases where the two coefficients are similar in magnitude but opposite in sign, then this fiscal policy change would have no independent effect on the fed funds rate; the only effect would occur through the indirect effect of fiscal policy on the output gap. For shorter horizons, the magnitude of the coefficient on S_{t+k} is larger in absolute value than that of $HEB6_{t+k}$, which would imply a reduction of the fed funds rate when both surpluses increase by the same amount, so in this case the Fed eases policy in response to structural fiscal tightening.

These results open the door to the possibility that fiscal policy affects monetary policy independently of its effect on the output gap. However, we caution that these results are not robust. In particular, one might argue that $(S_{t+k} - HEB6_{t+k})$, the "cyclical" component of the fiscal surplus, is strongly correlated with the rate of unemployment. In additional tests (not reported here), we found that

1) neither S_{t+k} nor $HEB6_{t+k}$ were statistically significant unless both were included in the regression.

2) when they were statistically significant, the sum of their coefficients was never statistically significantly different from zero.

3) When we replaced the output gap in the regression with the unemployment rate, our fiscal variables were never significant.

IV. Patton-Timmermann Tests

Patton and Timmermann (2012) describe methods for evaluating forecasts at multiple horizons, as is the case for our Greenbook fiscal forecasts. They evaluate a number of different tests but focus particular attention on "optimal revision regression tests" in section 3.3 of their paper. We run those tests for each of the

four major fiscal forecast variables (surplus, expenditures, receipts, and HEB), finding some small support for bias in the forecasts, but not extensive bias.

The optimal revision regression test examines regressions of either a realized variable on forecasts at different horizons, or regressing the last forecast made on forecasts made at longer horizons. The null hypothesis of rational forecasts requires that the coefficient on the most recent forecast have a coefficient of 1, while the coefficients on the constant term and on earlier forecasts are zero. Table 6 shows the results of these tests, with the p-values in the table showing the joint test of the null hypothesis on the zeroes and ones discussed above.

Table 6—Patton-Timmermann Optimal Revision Regression Test

Dependent Variable	Meeting	HEB	Surplus	Expenditures	Receipts
Actual	First	0.057	0.108	0.057	0.514
Actual	Last	0.348	0.180	0.467	0.525
Last forecast	First	0.904	0.012	0.199	0.889
Last forecast	Last	0.541	0.308	0.191	0.486

Note: P-values are shown for the test of the null hypothesis that the first lagged forecast has a coefficient of unity and all the other coefficients on lagged forecasts, as well as the constant, are zero.
Estimation is by least squares.
First and *Last* refer to the timing of the FOMC meeting within the quarter.

REFERENCES

Aruoba, S. Boragan. 2008. "Data Revisions Are Not Well Behaved." *Journal of Money, Credit, and Banking*, 40: 319–340.

Coibion, Olivier, and Yuriy Gorodnichenko. 2012. "Why Are Target Interest Rate Changes so Persistent?" *American Economic Journal: Macroeconomics*, 126–162.

Patton, Andrew J., and Allan Timmermann. 2012. "Forecast Rationality Tests Based on Multi-Horizon Bounds." *Journal of Business and Economic Statistics*, 30: 1–17.

Romer, Christina D, and David H Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *The American Economic Review*, 94(4): 1055–1084.

Rudebusch, Glenn D. 2006. "Monetary Policy Inertia: Fact or Fiction?" *International Journal of Central Banking*, 2(4): 85–135.

Taylor, John B. 1993. "Discretion Versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy*, 39: 195–214.