Gauging Taylor rules: Do differences make a difference?

Severin Bernhard* Tommaso Mancini Griffoli†‡

October 2011; first version April 2010

Abstract

The normative literature on Taylor rules is vibrant with debates as to the value of one rule formulation versus another. But do these "differences make a difference" in practice, in terms of recommended interest rates when responding to actual data? And if differences do arise, which are the elements of rules most responsible for them? Their specification, their assumptions of parameter values, or their definitions of inflation and output gaps? This paper attempts to provide a systematic and rigorous analysis of these questions over the 1990 to 2010 sample period. While the normative literature mostly insists on specification and assumptions, it is the choice of data that makes the biggest difference to interest rates. These findings imply that debates as to the conduct of policy should primarily focus on data sources. Yet, even the most favorable choice of data does not satisfactorily explain the very loose policy of the Federal Reserve between 2002 and 2007.

JEL classification: TBD

Keywords: TBD

*University of Fribourg, severin.bernhard@unifr.ch

†Swiss National Bank, tommaso.mancini@snb.ch

‡Many thanks to Jean-Pierre Danthine for comments to have pushed us to explore the questions behind this paper, as well as to Katrin Assenmacher, Daniel Kaufmann, Bob King, Elmar Mertens, Michel Peytrignet, Marcel Savioz, John Taylor, Sébastien Wälti and Mark Watson for helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Swiss National Bank or University of Fribourg.

Simple monetary policy rules, or Taylor (1993) rules, seem to speak with a single voice. When inflation is above target, or output above potential, monetary policy should be tightened; simple enough. But as soon as one lends an attentive ear, the single voice gives way to a cacophonous rumble.

Indeed, no single Taylor rule stands out as the perfect rule in a literature vibrant with debates on optimal formulations. Some rules are preferred for their fit with the inflation-output volatility frontier. Some seem to generate higher welfare or best replicate the utility maximizing outcome. Yet others appear more robust to data or model uncertainty.

This paper begins with the normative literature. However, it views this literature with a practical eye, not with the goal of providing yet another judgement favoring one rule over another. This paper asks whether the differences among Taylor rules discussed in the normative literature really make a difference in practical terms, in so far as recommended interest rate paths. And if so, where does this difference come from? Specifically, which are the elements of rules which, if changed, most affect recommended interest rates? In other words, should debates on the reaction functions to be followed by central banks be as diverse as those on optimal Taylor rules? Or should those "practical" debates focus on just the choice of one or two key elements of rules – those that actually make a difference?

To answer these questions, we follow one approach with two variants. Our general approach is to specify a large set of Taylor rules taken directly from the normative literature, generate recommended interest rate paths for each rule given observed data and compare this output in a systematic fashion to the interest rate path stemming from a baseline rule. To make the comparison, we rely on charts, root mean squared errors and wherever possible more rigorous statistical methods. With these we gauge if the effects of changing specific elements of rules really make a large and significant difference to recommended interest rates. The two variants of this general approach concern the choice of baseline.

We begin by setting up a framework of analysis. We categorize Taylor rules according to three building blocks: specifications, assumptions and datasets. Specifications refer to rules responding to real-time versus realized data, or to contemporaneous, nowcast or forecasts of data. Assumptions cover the constants in the rules: the natural interest rate, the inflation target and the sensitivity parameters. And datasets encompass the various sources and definitions of inflation and output gaps over the 1990 to 2010 period. By combining one element from each building block, we are able to generate a vast number of possible rule formulations which we study sequentially.

In the first variant, we define an ad-hoc rule of known parametric form. We then vary one element of that rule at a time and measure deviations of resulting interest rate paths. This allows for a controlled experiment yet perhaps one somewhat removed from reality.

To fill the gap we adopt the second variant in which the baseline becomes the actual Federal Funs rate; observed but of unknown parametrization. We thus gauge which rule best fits these data, then incrementally alter the formulation of that best fitting rule. Using Giacomini and White (2006) tests of predictive ability we study which changes to which elements of that rule significantly worsen its fit. We do so much along the lines of Faust and Wright (2011).

This paper finds that changing Taylor rule formulations can have a significant effect on recommended interest rate paths. Yet, while normative debates focus on varying specifications and assumptions, the difference in interest rate paths mostly comes from varying data sources for inflation and to a lesser extent the output gap. All other changes to Taylor rules have an order of magnitude smaller impact on interest rate paths.

As for which rule best explains the Federal Funds rate, we find similar results. The choice of appropriate data brings the rules to their "best fitting frontier". Further changes to rule specifications or assumptions do not significantly improve, nor detract from, model fit.

This finding has implications for the recent debate between Taylor (2010) and Bernanke (2010) on the monetary policy of the Federal Reserve during the 2002-2007 years. While Taylor proposed a rule responding contemporaneously to realized CPI inflation and Bernarnke argued for a rule including expectations of real-time PCE inflation, it was mostly their respective picks of inflation – CPI versus PCE – that accounted for the difference in the interest rate paths each held up. But even to the extent that the two might have agreed on a common definition of inflation, perhaps even the one allowing for the closest fit with the observed Federal Funds rate, we suggest that deviations from the actual policy rate were greater in the contentious 2002-2007 sample than they were between 1990 and 2002. Thus, the Federal Reserve's very expansive policy during the pre-crisis period cannot be easily reconciled with a Taylor rule drawn from the normative literature.

1 Rules, from the literature to this paper

The various Taylor rules considered in the normative literature can be generated by varying time subscripts and parameters in the following general formulation:

$$\begin{aligned} i_t = r + \mathbb{E}_s[\pi_{t-1+k}] + \gamma_1 \left(\mathbb{E}_s[\pi_{t-1+k}] - \pi^*\right) \\ + \gamma_2 \left(\mathbb{E}_s[y_{t-1+\eta}] - \mathbb{E}_s[y^*_{t-1+\eta}]\right) \end{aligned} \tag{1}$$

where i_t is the recommended nominal interest rate to be set by the central bank – the Federal Funds rate in the case of the US – r is the average or steady state natural interest rate (equal to the real interest rate at steady state) usually thought to correspond to the natural rate of growth (alternatively, the growth rate of potential output), π_t is inflation, y_t is output, y^*_t is the natural output level (alternatively potential output), and thus $(y_t - y^*_t)$ is the output gap. Moreover, $\mathbb{E}_s$ is the expectation operator conditional on information known up to and including time $s \leq t$. Finally, γ_1 and γ_2 are the sensitivity parameters capturing the degree to which the central bank responds to either inflation or output deviations.

The above general formulation has three building blocks: specifications, assumptions and data sources. Individual rules can be constructed by combining elements from each block. We extract from the normative literature five different specifications, three assumptions, and twelve data sources. We then consider all relevant combinations of these. Details and insights on the subsequent discussion of normative Taylor rules can be found in Taylor and Williams (2010).

First, the debate between real time and realized data implies two specifications. The real time specification suggests using data as it was known in the past. This is as in Orphanides (2001 and 2003a), and indeed most papers thereafter. The realized specification instead favors using the last dataset available, including all data revisions. This is an in Molodstova and Papell (2009), though in that paper data is filtered on a rolling basis.

In equation (1), realized data corresponds to $s = T$, where T is today. Vintage or real time data instead implies $s = t - 1$, suggesting that in any given quarter t the central bank only fully knows information up to and including prior quarters.

The second debate on rule specifications involves the timing of variables: should the central bank react to contemporaneous, nowcasts or forecasts of variables? Rudebusch and Svensson (1999), Clarida, Gali, and Gertler

(2000) and Levin, Wieland, and Williams (2003) find that rules responding to inflation forecasts slightly improve optimality to the extent that these are for less than a year's horizon. Yet, Taylor (1999) argues that contemporaneous rules are more appropriate since they reflect the information unequivocally available to policy makers at any given time, and current inflation is anyways a forward looking variable. Batini and Nelson (2001) falls somewhere in the middle, arguing that the extent to which a rule should be forward looking depends on how expectations are formed.

In equation (1) contemporaneous rules correspond to $k = \eta = 0$, nowcast rules (whereby the central bank estimates the current quarter's data based on monthly data or flash estimates) to $k = \eta = 1$ and rules responding to inflation forecasts imply $k = 4$ and $\eta = 0$. We call these last two rule specifications forward looking rules.

We do not incorporate the debate on which variables other than the inflation and output gap should enter the specification of Taylor rules. This is for two reasons. First, because there is in fact little debate; the consensus is that adding further variables to those of equation (1) does not improve optimality. And second, we aim to avoid a potential laundry list of additional variables in order to ensure as systematic an analysis as possible.[1]

Taylor rules discussed in the normative literature also vary according to assumptions. We focus on three. First, significant discrepancy exists over the estimation of a country's natural interest rate (r in equation 1). Laubach and Williams (2003) argue this uncertainty should translate into softer or more prudent monetary policy responses. In this paper, we therefore consider $r \in [0.5, 5]\%$ in steps of 0.5pp. As we discuss later, this can also be interpreted as changes to the target inflation rate.

The further two assumptions regard the sensitivity parameters. Taylor (1993) recommends responding with similar strength to deviations from target inflation and the output gap. Woodford (2003) popularized the "Taylor

[1]Simple Taylor rules, in the spirit of Bryant, Hooper, and Mann (1993), survive the tests of optimality and robustness to model uncertainty, as suggested in Levin, Wieland, and Williams (2003) and Orphanides and Williams (2006), among others. Moreover, Williams (2003) suggests that adopting a Taylor rule yields utility matching that from Ramsey optimal policy, in the sense of Giannoni and Woodford (2005), Svensson (2011 forthcoming) and Woodford (2011 forthcoming). Levin and Williams (2003) and Orphanides and Williams (2008) actually suggest that Taylor rules may prove better than Ramsey policy under model mis-specification. One potential addition is a lagged interest rate term to account for smoothing. Yet, Levin and Williams (2003), Taylor and Wieland (2009) and Isard, Laxton, and Eliasson (1999) suggest such rules are not robust to model specification, especially in a partially backward looking model or one with persistent policy errors.

principle" by which whatever the sensitivity parameters, the nominal interest rate should increase more than proportionally to higher inflation so as to increase real interest rates and stabilize inflation. Yet, Rudebusch (2001) argues that data uncertainty over the size of the output gap warrants lower response coefficients. On the contrary, Orphanides and Williams (2008) argue that more weight should be put on inflation stabilization so as to offset the lack of robustness to model uncertainty. And Reifschneider and Williams (2000) suggest a lower coefficient on the output gap when nearing the zero lower bound. We make room for these debates by considering values of 0.2, 0.5 and 0.8 for both sensitivity parameters γ_1 and γ_2.

Finally, sources of data and forecasts for the variables in Taylor rules can vary widely. We consider a wide set of data in order to span what is commonly found in the literature:[2]

- *Inflation measures*: (i) headline consumer price index (CPI), (ii) CPI core, (iii) headline personal consumption expenditures index (PCE), (iv) PCE core and (v) GDP deflator (core measures exclude the more volatile food and energy prices).[3]
- *Output gap measures*: (i) direct measure from Federal Reserve Board Greenbook (Greenbook), (ii) actual output minus potential output from the Congressional Budget Office (CBO) and (iii) from applying the Hodrick-Prescott (HP) filter to output series.
- *Sources of inflation forecasts* (for forward-looking models): (i) median of individual forecasts reported by the Survey of Professional Forecasters (SPF), (ii) Greenbook.

[2]For the output gap, the use of HP trends is prevalent. Yet Orphanides and van Norden (2002, 2005) and Orphanides (2003b, 2007) emphasize that filtered data amplifies the impact of revisions to real time data due to end of sample instability. Regarding inflation, most papers use year-on-year inflation measures following arguments in Levin, Wieland, and Williams (2003). As for inflation forecasts, there are fewer debates in the normative literature; positive papers have recently turned to the Survey of Professional Forecasters in order to avoid the five year data lag in the Greenbook, as in Molodstova, Nikolsko-Rzhevskyy, and Papell (2008) and Orphanides and Williams (2008).

[3]Note that the Federal Reserve changed its focus from headline CPI to headline PCE inflation in February 2000 and then to CPI core in July 2004 (as reported in Orphanides and Wieland (2008)).

2 Data

Data were obtained from two main sources: the Federal Reserve Bank of St. Louis (ALFRED, Archive Federal Reserve Economic Data) and the Federal Reserve Bank of Philadelphia (PHIL, Real-Time Data Center). These provided realized and real time data for both GDP and most inflation measures, as well as forecasts for inflation and output (Greenbook and SPF) and potential output (CBO). Table 1 provides more details about the corresponding source, frequency and horizon for each series we use.

The interest in also working with GDP forecasts was to obtain smoother end of sample estimates of potential output when using HP trends. In this context, we extended all GDP series to $(t-1+4)$ using forecasts, filtered the full series then cut off filtered values at $(t-1)$.

Data were collected in quarterly frequency, for both vintages (the date at which a series becomes known) and periods (the frequency of the data series). Vintages range from the first quarter of 1990 to the fourth quarter of 2010, although not all vintages are not available for all series as shown in Table 2. Note that price measures are available in monthly vintages and periods. For these, we took every third month's vintage and built quarterly periods by averaging data over three months. Also, CBO vintages are annual between 1991 and 1998, and semi-annual thereafter. Over intermediate vintages we simply duplicated the most recent vintage series. Finally, when working with Greenbook data, we only considered data from the second Greenbook published each quarter (from the second scheduled FOMC decision of each quarter). We did so for consistency, since SPF data is released at the end of the second month each quarter and all quarterly data incorporate information known by the end of the quarter.

3 Which elements of rules make a difference? (first variant)

Which elements of Taylor rules, be they part of specifications, assumptions or datasources, make the biggest difference to recommended interest rates? We answer the question by establishing a baseline rule, changing elements of that rule one at a time, then comparing interest rate paths using root mean squared errors (RMSEs) as well as charts.

Following standard methodology we compute the RMSE as:

$$RMSE_j = \left(\frac{1}{N} \sum_{t=0}^{N} (i_{j,t} - i_{b,t})^2 \right)^{1/2} \tag{2}$$

where j captures a given Taylor rule formulation, b captures the baseline formulation and N represents the length of the sample under investigation, with $N \leq T$ to allow for sub-samples.

The baseline rule we choose is a contemporaneous rule ($k = \eta = 0$), using real time data ($s = (t-1)$), with an average natural (real) interest rate (r) of 2.5[4], an inflation target (π^*) of 2[5] and inflation measured by headline CPI and the output gap as the difference between CBO potential output and actual GDP.

Results are presented in Table 3, divided into four panels representing contemporaneous, inflation forecasting and nowcast rules based on real-time data, as well as contemporaneous rules based on realized data. We consider two data samples: a sub-sample from 1990 to 2005 and the full sample from 1990 to 2010. The first covers the period during which Greenbook data is available.[6]

We realize, first, that data matters most to Taylor rule interest rates. Using measures of inflation other than CPI increases RMSEs substantially from the baseline (as seen in Table 3, Panel 1). The same is true when changing measures of the output gap, although corresponding RMSEs are somewhat smaller on average. Results are illustrated in Figures 1 and 2.

Second, much along the same lines, the difference between the contemporaneous baseline and forward looking specifications mostly comes from data. RMSEs increase noticeably when going from the baseline to forwarding looking rules (compare the first and second panels, or the first and third panels

[4]Taylor (1993) sets r to 2%, while the average CBO estimate of potential output from 2010 and 2007 vintages starting from 1990 is closer to 3%. We therefore take the average of the two. Most papers that estimate Taylor rules assume r is constant over time and subsume it in the intercept term, as in Orphanides and Williams (2008) and Orphanides and Wieland (2008).

[5]As suggested in the Federal Reserve Board minutes of February 1, 2005, regarding the discussion of "Considerations pertaining to the establishment of a specific numerical price related objective for monetary policy." The objective of 2% regards headline CPI inflation, while a target of 1.5% for PCE inflation was judged appropriate.

[6]Note that within any one panel we only present results that are directly comparable in the sense that they share at least 90% of the sample. Recall from the data section that some series do not span the entire sample, such as real-time core PCE data, or real-time forecasts of PCE core PCE and core CPI.

of Table 3). But the difference in RMSE mostly emerges when CPI inflation appears in the rules. The change in RMSE is instead slight if GDP deflators or CPI core inflation enter Taylor rules. The reason comes from CPI inflation being more volatile and thus difficult to forecast. When forecast errors are large, contemporaneous and forward looking rules most diverge. Inflation forecast errors are shown in Figure 3, while comparisons of Taylor rule specifications appear in Figure 4.

Third, adding to the above, the source of inflation forecasts makes little difference to forward looking models. Whether forecasts are taken from the SPF or Greenbook, RMSEs with contemporaneous rules are approximately equal on average. This is due in great part to inflation forecasts being rather similar across sources as shown in Figure 5.

Fourth, using realized instead of real-time data does alter the baseline interest rate path, but not by very much and mostly in specifications where it is least expected. Figure 6 contrasts interest rate paths for the baseline with the same rule based on realized data. The divergence is smaller than when changing data sources and appears only in certain periods. Those naturally coincide with the periods when output gap revisions were largest; these are shown in Figure 7 (recall only output gap data, and not inflation, get revised). Finally, comparing the top to the bottom panels of Table (3) suggests the difference between real time and realized data narrows when rules respond to HP output gaps as opposed to gaps computed by the CBO. This result is somewhat surprising in light of Orphanides and van Norden (2002, 2005) and Orphanides (2003b, 2007) which suggest that filtering techniques usually amplify the difference between realized and real time data. This may be due to CBO output gaps relying on models whose estimations of potential output are less sensitive to data revisions than the HP filter, and thus amplify fluctuations in the output gap when output gets revised.

Finally, changing assumptions does make a noticeable difference to Taylor rule interest rates, but in an uninteresting way from the standpoint of sensitivity analysis. Indeed, changing assumptions mostly shifts interest rate paths up or down to an arbitrarily large extent. The baseline Taylor rule can be re-written as $i_t = C + (1+\gamma_1)\pi_{t-1} + \gamma_2 \left(y_{t-1} - y^*_{t-1}\right)$ where $C \equiv (r - \gamma_1 \pi^*)$ is a constant including three of the four variables affected by assumptions. The implications of changing this constant are shown in Figure 8 and the very similar effects of changing sensitivity parameters alone are shown in Figures 9 and 10. We return to the effects of assumptions in the next section whose more realistic setting provides bounds on plausible parameter values.

4 Taylor rules and the Federal Funds rate, the devil is in the data (second variant)

This section provides further tests of the above takeaways in a more realistic setting and using a more rigorous approach. The baseline here becomes the Federal Funds rate. Yet, because we have no explicit parametrization of that series we begin by gauging which of the Taylor rules under consideration yields the closest interest rate path in terms of RMSE. We then change each element of that best fitting rule and generating corresponding interest rate paths, much as we did earlier. This time, though, we judge the statistical difference between these paths in their ability to forecast the Federal Funds rate one quarter ahead. The elements which significantly worsen the best rule's fit are deemed to be those that make the biggest difference.

We judge the difference in two rules' ability to explain the Federal Funds rate with the Giacomini and White (2006) test of predictive ability based on the null hypothesis:

$$H_0: \quad \mathbb{E}_t \left[L_{t+\tau}(Y_{t+\tau}, \widehat{f}_t) - L_{t+\tau}(Y_{t+\tau}, \widehat{g}_t) \right] = 0 \qquad (3)$$

where L is a loss function, in our case the squared error, of the forecast of the effective Federal Fund rate ($Y_{t+\tau}$), given a model (either rule $\widehat{f}$ or $\widehat{g}$). We work with $\tau = 1$ and the unconditional version of the test, as we are interested in which rule was more accurate, on average, in the past instead of which forecast will be more accurate for a specific future date, as would be the case for the conditional test. If the statistic is positive, rule $\widehat{g}$ is deemed to have a closer fit to the data, while $\widehat{f}$ is preferred in the case of a negative statistic.[7]

First, we find that the best-fitting rule for the 1990-2002 sample as well as the full 1990-2010 sample is the nowcast rule based on real time CBO potential output and GDP deflators (see Figure 11). In this finding, we rejoin Orphanides and Wieland (2008) which suggests rules responding to short term inflation forecasts better fit the Federal Funds rate. RMSEs of this best-fitting are presented in the third panel of Table 4.

Second, it is the appropriate choice of data which brings rules to their "best fitting frontier" (the set of statistically indistinguishable rules in terms of fit with the Federal Funds rate). A glance across panels 1, 2 and 3 of Table

[7] The Giacomini and White (2006) test statistic is shown to have a t-distribution. The test statistic found with a suitable HAC estimator of the asymptotic variance of the difference in loss functions coincides with that proposed by Diebold and Mariano (1995).

4 (corresponding once again to the contemporaneous, inflation forecast and nowcast rules, respectively), shows that it is the use of GDP deflators and CBO output gaps which always yields lowest RMSEs.

Giacomini and White (2006) test results confirm this finding. As the top panel of Table 5 suggests, the test statistic comparing the CPI and GDP deflator-based rules is significant at the 5% level in all subsamples. Instead, the middle panel of Table 5 shows that the contemporaneous and forward looking rules based on GDP deflators and CBO output gaps are not significantly different from one another. Likewise, the source of forecasts for forward looking and nowcast rules also makes no significant difference to the best fitting rules (results not shown).

These results continue to hold in the 2002-2007 subsample. Here, the consistent availability of a wider set of data allows us to find another rule formulation with slightly better RMSE (based on PCE core inflation in the realm of contemporaneous rule specifications). Yet, this rule is not significantly different from the contemporaneous or nowcast rules based on GDP deflators.

Third, changing assumptions has the same effect of changing rule specifications: RMSEs improve, but not in a statistically significant way. Over the 1990-2002 sub-sample, the standard assumptions (of 2.5% for the natural real interest rate, 2% for target inflation and 0.5 for both sensitivity parameters γ_1 and γ_2) yield the best RMSE, as shown in Table 6. Over the full sample, there is some room for improvement by assuming a 2% real interest rate (alternatively, a target inflation rate of 3%), yet the difference in RMSE is not significant (see panel 3 of Table 5). Over the middle sample, two adjustments to assumptions improve the RMSE of the best-fitting rule: a 2% real interest rate and a 0.8 coefficient on the inflation gap. The first could be justified by the post internet bubble and the second by greater model uncertainty or nearing the zero lower bound, as in Reifschneider and Williams (2000). But once again, the difference in RMSE is not significant. In the end, Taylor (1993)'s original suggestion of sensitivity parameters equally set to 0.5 continues to best describe the data over all sub-samples.

Finally, our findings suggest that debates on the conduct of policy should primarily focus on the choice of data. Our results imply that the interest rate paths produced by Taylor (2010) and Bernanke (2010), with rules responding to contemporaneous realized CPI inflation versus expected real-time PCE inflation, mostly diverged because of their their different choices of inflation. Yet even the inflation measure which best first the Federal Funds rate in the contentious 2002-2007 period does not offer a very close fit (see Figure

12). While it is not possible to compare rules across different sub-samples with statistical techniques, it is striking that the rule with the lowest RMSE in the 2002-2007 subsample still exhibits an error noticeably larger than its counterpart in the pre-2002 sample or the full sample. The difference is of 70 and 35%, respectively, with standard assumptions and 25 and 10% with optimized assumptions. In fact, RMSEs increase in the 2002-2007 subsample relative to the earlier sample for nearly all possible rule formulations.

5 Conclusion

This paper set out to view the normative literature with a practical eye. The paper's findings should therefore have some appeal to policy makers involved in the practical implementation of Taylor rules. The practical implications of this paper's findings would be:

First, policy makers should be aware or weary of data. The choice of which measure of inflation and to a lesser extent output gap to include in one's rule will significantly affect recommended interest rates. It may be wise to hedge by considering various data sources.

A second hedge could be to give more weight to a smooth measure of inflation, such as core CPI or GDP deflators. This will dampen the difference between choosing a contemporaneous or forward-looking rule, based on inflation forecasts.

Third, in that case, policy makers may as well follow a contemporaneous rule for simplicity, as data does not have to be forecasted.

Fourth, in specifying a rule, Taylor (1993)'s original proposal of responding with equal sensitivity to deviations from target inflation and the output gap is a good starting point. Indeed, the difference in interest rates generated by changing these assumptions does not tend to be significant, at least not in the case of the US.

And finally, from the hypothetical to the concrete, the interest rate policy of the Federal Reserve between 2002 and 2007 seems to have been overly accommodative relative to even the best fitting Taylor rules.

References

Batini, N., and E. Nelson (2001): "Optimal horizons for inflation targeting," *Journal of Economic Dynamics and Control*, 25(6-7), 891–910.

Bernanke, B. S. (2010): "Monetary policy and the housing bubble," *Speech at the Annual Meeting of the American Economic Association, Atlanta Georgia.*

Bryant, R., P. Hooper, and C. Mann (1993): *Evaluating policy regimes: new empirical research in empirical macroeconomics*. Brookings Institution.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary policy rules and macroeconomic stability: evidence and some theory," *Quarterly Journal of Economics*, 115(1), 147–180.

Diebold, F. X., and R. S. Mariano (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Faust, J., and J. Wright (2011): "Forecasting inflation," *Handbook of forecasting*, Forthcoming.

Giacomini, R., and H. White (2006): "Tests of conditional predictive ability," *Econometrica*, 74(6), 1545–1578.

Giannoni, M. P., and M. Woodford (2005): "Optimal inflation targeting rules," in *The inflation targeting debate*, ed. by B. S. Bernanke, and M. Woodford, pp. 93–162. University of Chicago Press.

Isard, P., D. Laxton, and A.-C. Eliasson (1999): "Simple monetary policy rules under model uncertainty," *International Monetary Fund Working Paper no. 99/75.*

Laubach, T., and J. C. Williams (2003): "Measuring the natural rate of interest," *Review of Economics and Statistics*, 85(4), 1063–1070.

Levin, A., V. Wieland, and J. C. Williams (2003): "The performance of forecast-based monetary policy rules under model uncertainty," *American Economic Review*, 93(3), 622–645.

Levin, A. T., and J. C. Williams (2003): "Robust monetary policy with competing reference models," *Journal of Monetary Economics*, 50, 945–975.

Molodstova, T., A. Nikolsko-Rzhevskyy, and D. H. Papell (2008): "Taylor rules with real-time data: a tale of two countries and one exchange rate," *Journal of Monetary Economics*, 55(Supplement), S63–S79.

Molodstova, T., and D. H. Papell (2009): "Out-of-sample exchange rate predictability with Taylor rule fundamentals," *Journal of International Economics*, 77(2), 167–180.

Orphanides, A. (2001): "Monetary policy rules based on real time data," *American Economic Review*, 94(4), 964–985.

——— (2003a): "Historical monetary policy analysis and the Taylor rule," *Journal of Monetary Economics*, 50(5), 983–1022.

——— (2003b): "Monetary policy evaluation with noisy information," *Journal of Monetary Economics*, 50(3), 605–631.

——— (2007): "Taylor Rules," *Mimeo.*

Orphanides, A., and S. van Norden (2002): "The unreliability of output gap estimates in real time," *Review of Economics and Statistics*, 84(4), 569–583.

——— (2005): "The reliability of intlation forecasts based on output gap estimates in real time," *Journal of Money, Credit and Banking*, 37(3), 583–601.

Orphanides, A., and V. Wieland (2008): "Economic projections and rules of thumb for monetary policy," *Federal Reserve Bank of St. Louis Review*, 90(4), 307–324.

Orphanides, A., and J. C. Williams (2006): "Monetary policy with imperfect knowledge," *Journal of the European Economic Association*, 4(2-3), 366–375.

——— (2008): "Learning, expectations formation and the pitfalls of optimal control monetary policy," *Journal of Monetary Economics*, 55S, S80–S96.

Reifschneider, D. L., and J. C. Williams (2000): "Three lessons for monetary policy in a low inflation era," *Journal of Money, Credit and Banking*, 32(4), 936–966.

Rudebusch, G. D. (2001): "Is the Fed too timid? Monetary policy in an uncertain world," *Review of Economics and Statistics*, 83, 203–217.

Rudebusch, G. D., and L. E. O. Svensson (1999): *Policy rules for inflation targeting*pp. 203–253. University of Chicago Press.

Svensson, L. E. O. (2011 forthcoming): "Inflation targeting," in *Handbook of monetary economics*. Elsevier.

Taylor, J. B. (1993): "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

——— (1999): "The robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the European Central Bank," *Journal of Monetary Economics*, 43(3), 655–679.

——— (2010): "The Fed and the crisis: a reply to Ben Bernanke," *The Wall Street Journal*, p. A19.

Taylor, J. B., and V. Wieland (2009): "Surprising comparative properties of monetary models: results from a new database," *NBER Working Paper 14849*.

Taylor, J. B., and J. C. Williams (2010): "Simple and robust rules for monetary policy," *Working paper, Federal Reserve Bank of San Francisco*, 10.

Williams, J. C. (2003): "Simple rules for monetary policy," *Federal Reserve Bank of San Francisco Economic Review*.

Woodford, M. (2003): *Interest and prices: foundations of a theory of monetary policy*. Princeton University Press, Princeton, NJ.

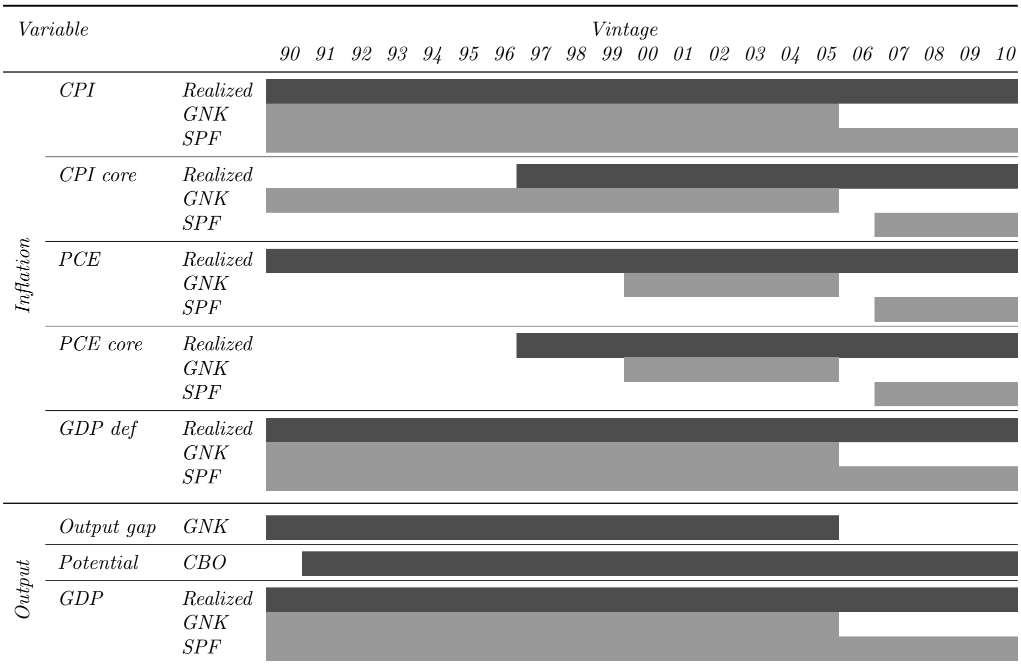
——— (2011 forthcoming): "Optimal monetary stabilization policy," in *Handbook of monetary economics*. Elsevier.

Table 1: *Data for inflation and output gap: summary (** note on Frequency Vint. and Per.)*

	Variable		*Source*	*Frequency*		*Vintages*	
				Vint.	*Per.*	*Start*	*End*
Inflation	*CPI*	*realized*	PHIL	Q	M	1972q3	2010q4
		GNK	ALFRED	8/y	Q	1979q4	2005q4
		SPF	PHIL	Q	Q	1981q3	2010q4
	CPI core	*realized*	ALFRED	M	M	1996q1	2010q4
		GNK	ALFRED	8/y	Q	1986q1	2005q4
		SPF	PHIL	Q	Q	2007q1	2010q4
	PCE	*realized*	PHIL	Q	Q	1965q4	2010q4
		GNK	ALFRED	8/y	Q	2000q1	2005q4
		SPF	PHIL	Q	Q	2007q1	2010q4
	PCE core	*realized*	PHIL	Q	Q	1996q1	2010q4
		GNK	ALFRED	8/y	Q	2000q1	2005q4
		SPF	PHIL	Q	Q	2007q1	2010q4
	GDP def	*realized*	PHIL	Q	Q	1965q4	2010q4
		GNK	ALFRED	8/y	Q	1986q1	2005q4
		SPF	PHIL	Q	Q	1968q4	2010q4
Output	*Output gap*	*GNK*	PHIL	8/y	Q	1987q3	2005q4
	Potential	*CBO*	ALFRED	1/y	Q	1991	1998
				2/y	Q	1999:1	2009:2
	GDP	*realized*	PHIL	Q	Q	1965q4	2010q4
		GNK	ALFRED	8/y	Q	1978q1	2005q4
		SPF	PHIL	Q	Q	1968q4	2010q4

Data for inflation and the output gap come from a variety of sources, depending on whether the data is realized (known as of today) or real-time/vintage data. Data frequency is mostly quarterly. Not all vintage series span the entire sample of interest, going from 1990 to 2010. ALFRED stands for Archive of Federal Reserve Economic Data, from the Federal Reserve Bank of St. Louis. And PHIL stands for the Real-Time Data Center of the Federal Reserve Bank of Philadelphia. GNK stands for the Federal Reserve Board's Greenbook. And SPF stands for the Survey of Professional Forecasters.

Table 2: *Data for inflation and output gap: availability*



Not all vintage data spans the entire sample of interest, going from 1990 to 2010. See prior table for an explanation of the various acronyms.

TABLE 3: *RMSE relative to baseline*

PANEL 1: Contemporaneous, real time data rules

Sample	*1990-2005*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
Inflation								
CPI	0.41	*base*	0.99	0.99		*base*	1.19	
CPI core	1.17	1.03	1.52	1.52		1.46	1.96	
PCE	1.11	0.91	1.41	1.42		0.90	1.55	
PCE core								
GDP def	1.34	1.17	1.52	1.54		1.47	1.87	

PANEL 2: Forward looking, real time data rules

Sample	*1990-2005*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
SPF inflation								
CPI	1.08	0.99	1.31	1.33		1.50	1.91	
CPI core								
PCE								
PCE core								
GDP def	1.27	1.13	1.41	1.43		1.46	1.83	
GNK inflation								
CPI	1.37	1.25	1.51	1.52				
CPI core	1.07	0.95	1.27	1.28				
PCE								
PCE core								
GDP def	1.42	1.26	1.54	1.56				

PANEL 3: Nowcast, real time data rules

Sample	*1990-2005*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
SPF inflation								
CPI	0.72	0.54	1.11	1.14		0.97	1.48	
CPI core								
PCE								
PCE core								
GDP def	1.40	1.21	1.56	1.58		1.51	1.91	
GNK inflation								
CPI	0.83	0.72	1.19	1.21				
CPI core	1.15	1.00	1.48	1.52				
PCE								
PCE core								
GDP def	1.36	1.17	1.52	1.54				

PANEL 4: Contemporaneous, realized data rules

Sample	*1990-2005*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
Inflation								
CPI		0.64	0.99			0.64	1.22	
CPI core		1.14	1.48			1.43	1.85	
PCE		1.18	1.42			1.07	1.53	
PCE core		1.43	1.69			1.57	1.96	
GDP def		1.36	1.53			1.56	1.80	

Root mean squared errors relative to baseline. No results are reported for those formulations with insufficient observations to warrant a meaningful comparison to others. Note, GNK = Federal Reserve Board's Greenbook, CBO = Congressional Budget Office, SPF = Survey of Professional Forecasters, HP = Hodrick-Prescott filter, based on real-time series plus real-time forecasts coming from either SPF or GNK, CPI = Consumer Price Inflation, PCE = Personal Consumer Expenditures, GDP defl = GDP deflator, ex-oil = core inflation measure, excluding oil products.

Table 4: *RMSE relative to Federal Funds rate*

PANEL 1: Contemporaneous, real time data rules

Sample	*1990-2002*				*2002-2007*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
Inflation												
CPI	1.44	1.51	1.78	1.78		2.16	2.64			2.16	2.58	
CPI core	1.61	1.65	2.06	2.06		1.78	2.02			1.57	2.18	
PCE	1.03	1.00	1.54	1.54		1.61	2.02			1.60	2.18	
PCE core						1.35	1.51					
GDP def	0.92	0.93	1.39	1.40		1.42	1.88			1.21	1.84	

PANEL 2: Forward looking, real time data rules

Sample	*1990-2002*				*2002-2007*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
SPF inflation												
CPI	1.37	1.44	1.70	1.71		2.02	2.36			1.57	2.12	
CPI core												
PCE												
PCE core												
GDP def	0.88	0.94	1.33	1.34		1.63	1.91			1.17	1.78	
GNK inflation												
CPI	1.22	1.26	1.58	1.58								
CPI core	1.32	1.38	1.67	1.67								
PCE												
PCE core												
GDP def	0.85	0.85	1.37	1.37								

PANEL 3: Nowcast, real time data rules

Sample	*1990-2002*				*2002-2007*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
SPF inflation												
CPI	1.29	1.34	1.65	1.66		2.25	2.74			2.02	2.44	
CPI core												
PCE												
PCE core												
GDP def	0.83	0.82	1.34	1.36		1.47	1.93			1.10	1.79	
GNK inflation												
CPI	1.26	1.32	1.60	1.61								
CPI core	1.43	1.50	1.91	1.94								
PCE												
PCE core												
GDP def	0.83	0.84	1.33	1.35								

PANEL 4: Contemporaneous, realized data rules

Sample	*1990-2002*				*2002-2007*				*1990-2010*			
Output gap	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}	*GNK*	*CBO*	HP_{SPF}	HP_{GNK}
Inflation												
CPI		1.66	1.83			2.55	2.56			2.37	2.65	
CPI core		1.79	2.04			1.93	1.91			1.72	2.16	
PCE		1.26	1.47			2.20	2.21			1.94	2.28	
PCE core		1.30	1.58			1.75	1.75			1.48	1.96	
GDP def		1.29	1.43			2.55	2.58			1.78	2.07	

Root mean squared errors relative to the Federal Funds rate. See Table 3 for details on acronyms.

Table 5: *Giacomini and White (2006) test results*

PANEL 1: Contemporary rule specification: baseline versus best fitting rule

Sample	*1990-2002*		*2002-2007*		*1990-2010*	
Rule formulation	*Base (1)*	*Rule (2)*	*Base (1)*	*Rule (2)*	*Base (1)*	*Rule (2)*
(1) baseline rule						
Test statistic	-	2.18	-	3.55	-	3.14
(p-value)		*(0.035)*		*(0.002)*		*(0.002)*
(2) Contemp/GDP defl/CBO						
Test statistic	-2.18	-	-3.55	-	-3.14	-
(p-value)	*(0.035)*		*(0.002)*		*(0.002)*	

PANEL 2: Best fitting rules across different specifications

Sample	*1990-2002*			*2002-2007*			*1990-2010*		
Rule formulation	*Rule (1)*	*Rule (2)*	*Rule (3)*	*Rule (1)*	*Rule (2)*	*Rule (3)*	*Rule (1)*	*Rule (2)*	*Rule (3)*
(1) Contemp/GDP defl/CBO									
Test statistic	-	-0.02	1.29	-	-1.23	-0.67	-	0.42	1.52
(p-value)		*(0.982)*	*(0.203)*		*(0.233)*	*(0.510)*		*(0.676)*	*(0.133)*
(2) Fwd/GDP defl/CBO									
Test statistic	0.02	-	0.77	1.23	-	0.93	-0.42	-	0.78
(p-value)	*(0.982)*		*(0.445)*	*(0.233)*		*(0.363)*	*(0.676)*		*(0.437)*
(3) Nowcast/GDP defl/CBO									
Test statistic	-1.29	-0.77	-	0.67	-0.93	-	-1.52	-0.78	-
(p-value)	*(0.203)*	*(0.445)*		*(0.510)*	*(0.363)*		*(0.133)*	*(0.437)*	

PANEL 3: Best fitting rules and optimized assumptions

Sample	*1990-2002*		*2002-2007*		*1990-2010*	
Rule formulation	*Rule (1)*	*Rule (2)*	*Rule (1)*	*Rule (2)*	*Rule (1)*	*Rule (2)*
*(1) Best fitting rule (BFR) **						
Test statistic	-	0.00	-	0.82	-	0.69
(p-value)				(0.418)		(0.495)
(2) BFR, optimized assumptions						
Test statistic	0.00	-	-0.82	-	-0.69	-
(p-value)			(0.418)		(0.495)	

** Nowcast/GDPdefl/CBO for 1990-2002 and 1990-2010; Contemp/PCEcore/CBO for 2002-2007*

Giacomini and White (2006) test results, value of test statistic and p-value. Tests compare the statistical significance of the difference between the rule indicated in the row with that indicated in the column, in the ability to forecast the Federal Funds rate one quarter ahead. When the test statistic is positive, the rule indicated in the column is preferred to that in the row. The p-value indicates the significance of the difference in the two rules.

Table 6: *RMSE for different assumptions*

Panel 1: Best rule sample 1990-2002: Nowcast, GDP deflator, CBO output gap

Sample						*1990-2002*					
Real interest rate		*0.5*	*1.0*	*1.5*	*2.0*	*2.5*	*3.0*	*3.5*	*4.0*	*4.5*	*5.0*
γ_1	γ_2										
0.2	*0.2*	2.254	1.838	1.476	1.216	1.13	1.255	1.541	1.917	2.339	2.788
	0.5	2.139	1.69	1.279	0.957	0.834	0.988	1.325	1.743	2.195	2.664
	0.8	2.311	1.897	1.537	1.274	1.177	1.283	1.552	1.916	2.331	2.774
0.5	*0.2*	2.269	1.863	1.515	1.273	1.203	1.331	1.611	1.98	2.397	2.841
	0.5	2.107	1.657	1.246	0.927	0.816	0.986	1.333	1.756	2.211	2.683
	0.8	2.237	1.814	1.442	1.169	1.075	1.201	1.493	1.875	2.303	2.755
0.8	*0.2*	2.302	1.911	1.581	1.361	1.305	1.433	1.704	2.063	2.471	2.908
	0.5	2.096	1.651	1.248	0.943	0.849	1.026	1.373	1.794	2.247	2.717
	0.8	2.181	1.752	1.373	1.094	1.005	1.151	1.462	1.858	2.294	2.753

Panel 2: Best rule sample 2002-2007: Contemporaneous, PCE core, CBO output gap

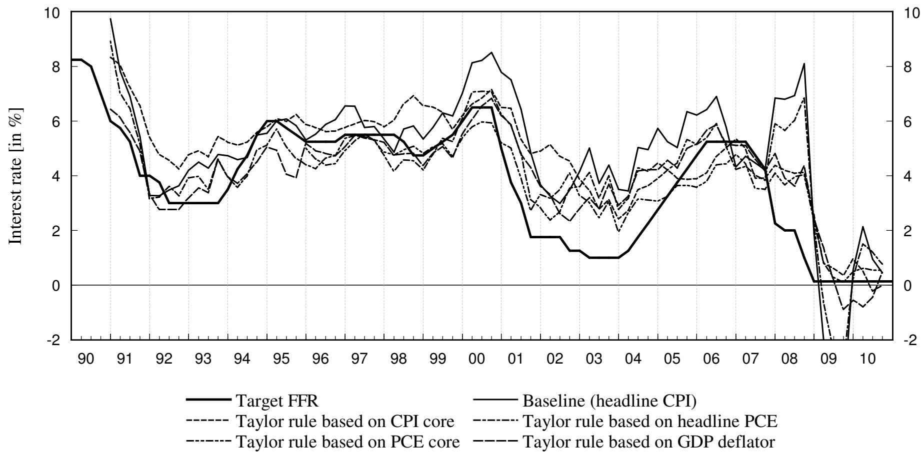
Sample						*2002-2007*					
Real interest rate		*0.5*	*1.0*	*1.5*	*2.0*	*2.5*	*3.0*	*3.5*	*4.0*	*4.5*	*5.0*
γ_1	γ_2										
0.2	*0.2*	1.693	1.436	1.326	1.397	1.627	1.96	2.353	2.781	3.229	3.691
	0.5	1.948	1.62	1.398	1.335	1.454	1.716	2.067	2.47	2.904	3.356
	0.8	2.242	1.868	1.567	1.385	1.373	1.532	1.82	2.186	2.597	3.034
0.5	*0.2*	1.694	1.405	1.258	1.3	1.515	1.845	2.238	2.668	3.118	3.582
	0.5	1.966	1.614	1.359	1.261	1.353	1.604	1.953	2.357	2.792	3.246
	0.8	2.273	1.882	1.555	1.34	1.293	1.431	1.71	2.074	2.485	2.925
0.8	*0.2*	1.706	1.389	1.203	1.211	1.409	1.733	2.127	2.558	3.01	3.476
	0.5	1.994	1.622	1.335	1.199	1.261	1.498	1.843	2.247	2.684	3.14
	0.8	2.313	1.907	1.557	1.309	1.225	1.338	1.605	1.966	2.378	2.818

Panel 3: Best rule full sample: Nowcast, GDP deflator, CBO output gap

Sample						*1990-2010*					
Real interest rate		*0.5*	*1.0*	*1.5*	*2.0*	*2.5*	*3.0*	*3.5*	*4.0*	*4.5*	*5.0*
γ_1	γ_2										
0.2	*0.2*	1.89	1.59	1.409	1.393	1.548	1.831	2.193	2.601	3.037	3.49
	0.5	1.927	1.529	1.21	1.043	1.101	1.355	1.721	2.141	2.59	3.055
	0.8	2.339	1.941	1.599	1.361	1.282	1.392	1.652	2.005	2.411	2.847
0.5	*0.2*	1.885	1.588	1.412	1.402	1.561	1.846	2.209	2.617	3.053	3.506
	0.5	1.909	1.511	1.193	1.031	1.096	1.357	1.727	2.15	2.6	3.066
	0.8	2.313	1.913	1.571	1.333	1.259	1.376	1.643	2.002	2.412	2.85
0.8	*0.2*	1.901	1.613	1.445	1.44	1.6	1.883	2.244	2.65	3.083	3.535
	0.5	1.912	1.52	1.211	1.059	1.13	1.39	1.757	2.178	2.626	3.091
	0.8	2.305	1.907	1.569	1.336	1.268	1.39	1.66	2.02	2.429	2.868

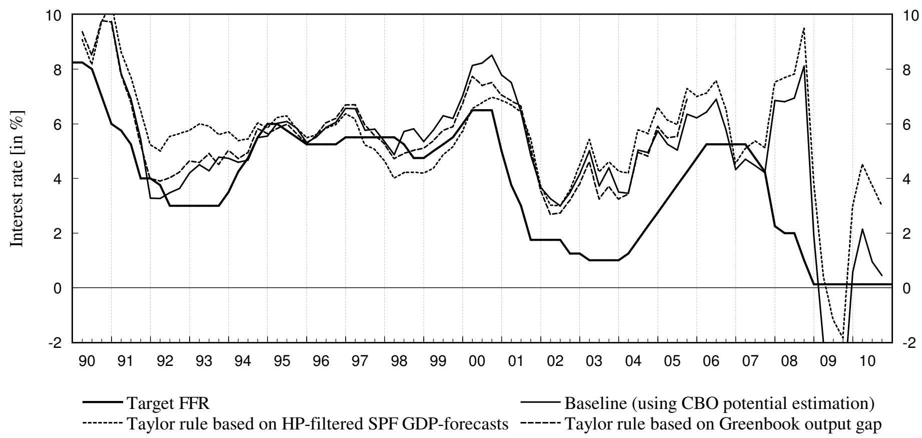
Root mean squared errors corresponding to different assumptions on the real interest rate and sensitivity parameters of the best fitting rules in each sub-sample (nowcast, GDP deflator, CBO output gap for the first and full sub-samples and contemporaneous, PCE core, CBO output gap for the middle sub-sample). Lighter cells indicate lower RMSEs.

Figure 1: *Effect of changing inflation measures*



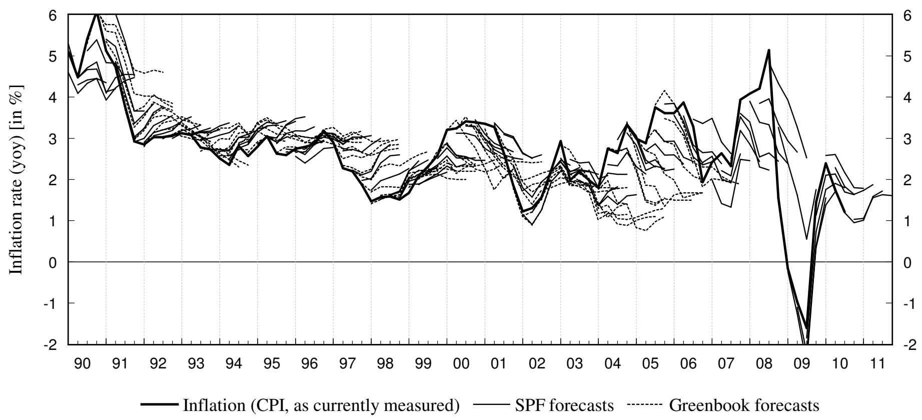
The effects of changing inflation measures on Taylor interest rates are large.

Figure 2: *Effect of changing output gap measure*



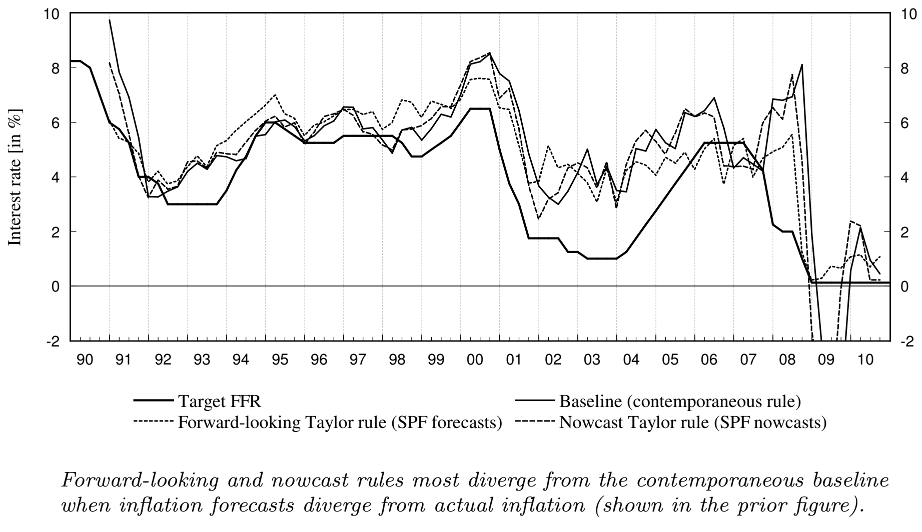
The effects of changing output gap measures are more noticeable in the 1990s than in the 2000s.

FIGURE 3: *Inflation forecasts*



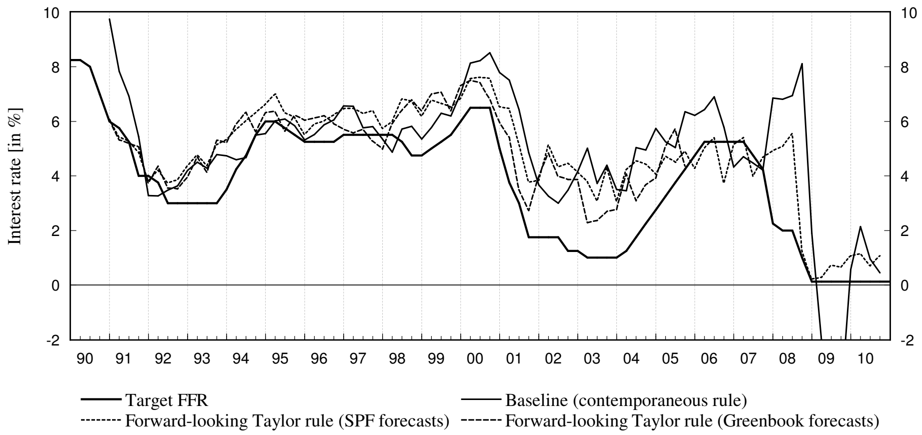
Four quarters ahead inflation forecasts from both the SPF and Greenbook are good when inflation is steady. But forecasts are rather poor when inflation volatility increases, such as in 1993-94, 1997-98, 2000-02, 2004-06 and again in 2008. It is in these same periods, naturally, that Taylor interest rates from forward-looking and nowcast rules most diverge from the baseline contemporaneous specification. At least, SPF and Greenbook forecasts are very similar during the sample in which they overlap.

FIGURE 4: *Effect of rule specification*



Forward-looking and nowcast rules most diverge from the contemporaneous baseline when inflation forecasts diverge from actual inflation (shown in the prior figure).

Figure 5: *Effect of inflation forecast source*

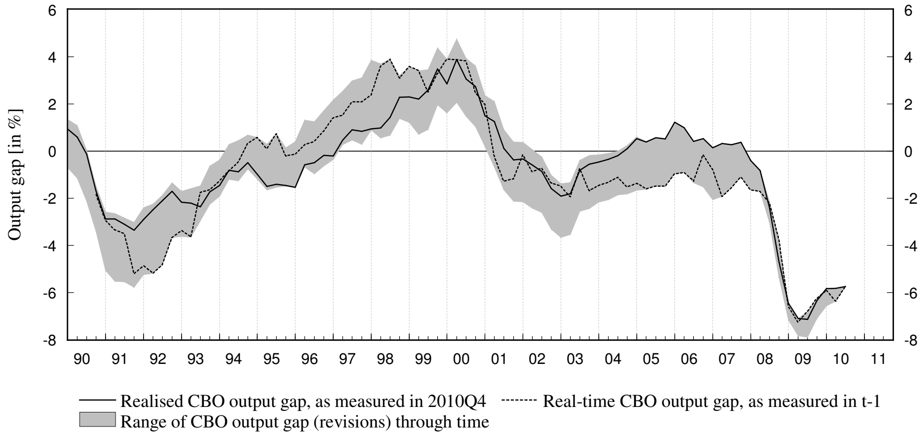


The effects of forecast source are neither very large nor systematic across the sample.

Figure 6: *Effect of information set*

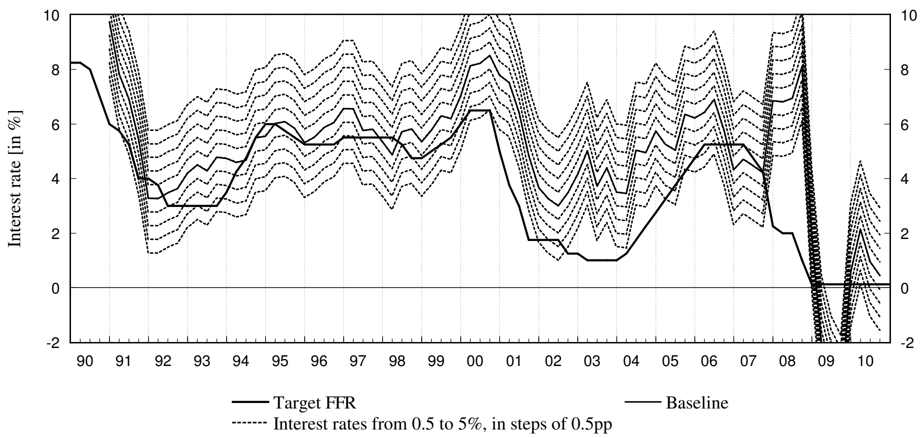
Using either real-time or realized data makes little difference throughout most of the sample.

Figure 7: *Revisions to the output gap*



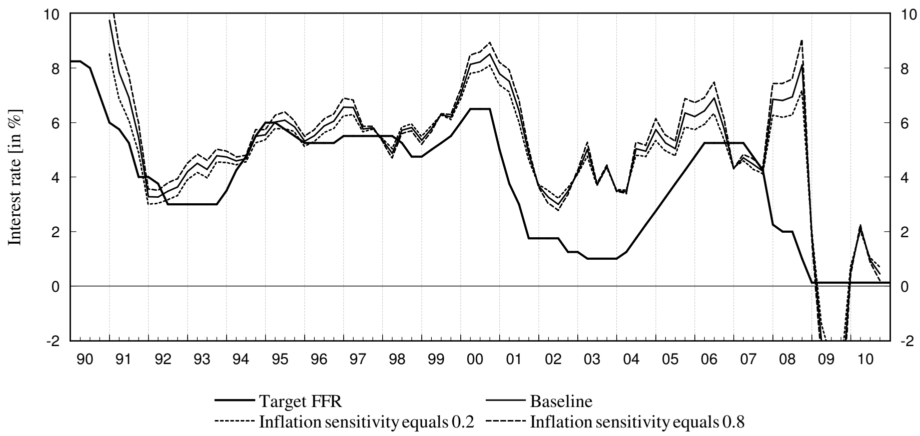
Real time CBO output gaps are shown with the dashed line. Realized gaps, instead, appear with the continuous line. The range of revisions having taken place in between is shown with the grey area. The size of the area comes from the fact that revisions are themselves revised, sometimes up, sometimes down. Until the late 1990s, revision of the output gap were generally downward. Then, after the turn of the millennium, output gaps were mostly revised up. In the most recent period, revisions are much smaller, but will grow as time passes. The ultimate revision, between the real time and realized datasets is the difference between the continuous and dotted lines. These were greatest from 1991-92, 1995-99 and 2004-07.

Figure 8: *Effect of changing natural interest rate*



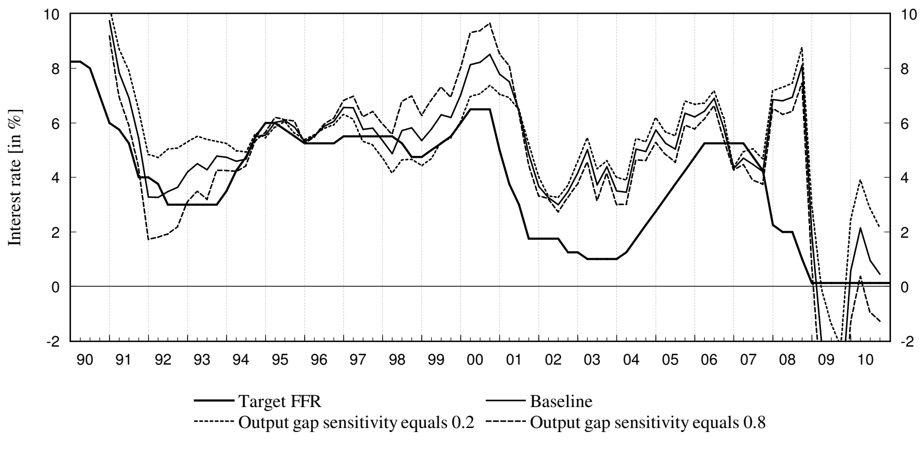
The effects of changing assumptions of the natural interest rate or the inflation target is akin to a level shift in Taylor Rule interest rates. The above captures, equivalently, a change in the natural interest rate between 0.5 *and* 5%*, or a target inflation rate between* 6 *and* 0% *and/or a mix thereof (which in many cases would be more realistic). The equivalence between these two sets of changes is captured in the text when re-writing the the Taylor rule after grouping constant terms.*

Figure 9: *Effect of changing inflation sensitivity*



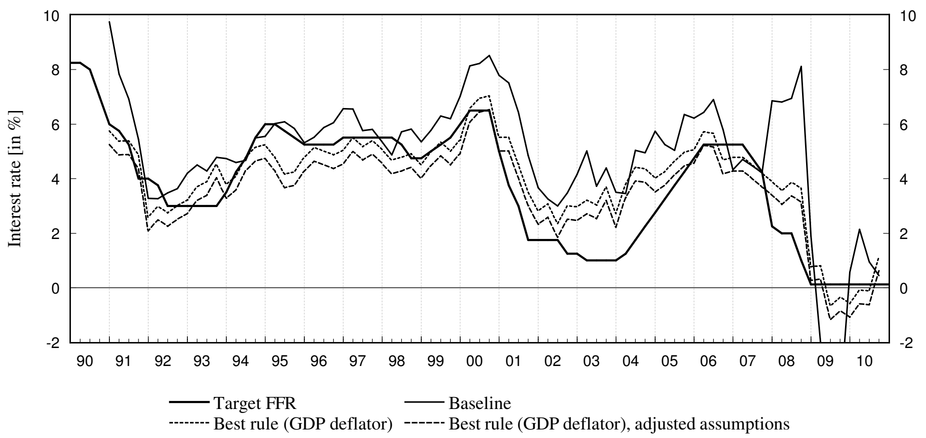
Changing policy makers' sensitivity to deviations of inflation from target has a relatively small effect on recommended interest rates.

Figure 10: *Effect of changing output gap sensitivity*



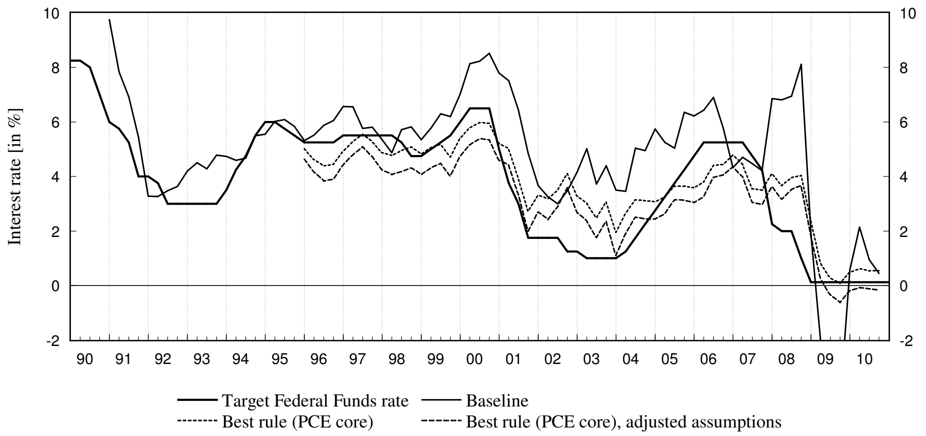
Changing policy makers' sensitivity to deviations of the output gap has a larger effect, especially in periods when such deviations are large.

Figure 11: *Best-fitting rule for the 1990-2002 and full sample*



The best-fitting rule for the 1990-2002 as well as full sample is the nowcast rule/GDP deflator/CBO output gap; the fit with the Federal Funds rate increases somewhat when assumptions are optimized to fit the full sample.

Figure 12: *Best-fitting rule for the 2002-2007 sample*



The best-fitting rule for the 2002-2007 sample is the contemporaneous rule/PCE core/ CBO output gap, although this rule is not statistically different from the best-fitting rule over the full sample (see prior figure), suggesting the Federal Reserve did not significantly change its reaction function in 2002. Root mean squared errors are larger on average than over the 1990-2002 sample even when assumptions are optimized to fit the 2002-2007 sample.