Real-time model uncertainty in the United States: the Fed from 1996-2003

Robert J. Tetlow and Brian Ironside*
Federal Reserve Board

First draft: September 22, 2004 This draft: February 10, 2005

Abstract

We study 30 vintages of FRB/US, the main macroeconomic model used by the Federal Reserve Board staff for forecasting and policy analysis since the model's inception in the mid-1990s. To do this, we exploit archives of the model code, coefficients, baseline databases and stochastic shock sets stored after each FOMC meeting over the period from July 1996 to November 2003. We document the changes in the model properties that occurred during this period–a surprisingly large and consequential set–and compute optimal Taylor-type rules for each vintage. The period of study was one of important changes in the U.S. economy with a productivity boom, a stock market boom and bust, a recession, the Asia crisis, the Russian debt default, corporate governance scandals and an abrupt change in fiscal policy. We compare these ex ante optimal rules against plausible alternatives. We find that model uncertainty is a substantial problem and that efficacy of purportedly optimal policy rules should not be taken on faith.

- **JEL Classifications:** E37, E5, C5, C6.
- **Keywords:** monetary policy, uncertainty, macro economic modeling, real-time analysis

---

* *Contact address:* Robert Tetlow, Federal Reserve Board, Washington, D.C. 20551. *Email:* rtetlow@frb.gov. The authors acknowledge the helpful comments of Richard Dennis, Karen Dury, Ed Nelson, Dave Reifschneider, Glenn Rudebusch, Simon van Norden, Peter von zur Muehlen, John C. Williams, Tony Yates, Fabrizio Zampolli and seminar participants at the Federal Reserve Board, the European Central Bank, the Bank of England, the FRB San Francisco and the Bank of Canada. We thank Flint Brayton for helping us interpret the origins of many of the model changes by vintage, and Douglas Battenberg for help getting the FRB/US model archives in working order. Part of this research was conducted while the first author was visiting the Bank of England and the FRBSF; he thanks those institutions for their hospitality. All remaining errors are ours. The views expressed in this paper are those of the authors alone and do not represent those of the Federal Reserve Board or other members of its staff. Find this paper and others at www.members.cox.net/btetlow/default.htm

1 Introduction

Over the past decade or so, there has been an explosion of work studying the characteristics of monetary policy rules in general, and interest-rate feedback rules in particular. Some papers have studied optimal targeting rules, in the parlance of Svensson (2002); other papers have studied instrument rules, sometimes with coefficients that have been chosen in an *ad hoc* fashion, and other times with coefficients that have been optimized.[1] The typical contribution in this area, posits a quadratic loss function,[2] a time-invariant, known linear model,[3] Gaussian disturbances, and a policy rule of a given form.[4] The optimal rule can then be constructed, sometimes analytically, but more often numerically, using established techniques. While considerable insight has come out of this literature, so has a fundamental critique, namely that rules formulated in this way may not be robust to misspecification. One of the earliest and most ardent critics of the standard practice was Bennett McCallum (1988). McCallum's specific proposal was to compare the candidate policies in a number of rival models and judge the "best" policy on its performance across the entire set. This argument has garnered a sizeable contingent of advocates in recent years with notable applications including Levin *et al.* (1999,2003). The rival models method has taken an important place alongside other methods, including analyses of parameter uncertainty and data uncertainty, as devices for "robustifying" monetary policy. [5]

Two other strands of the literature stand in contrast to these Bayesian approaches. The first of these is the robust control (or min-max) approach, championed in various

[1] Optimal targeting rules have been championed by Svensson (1999,2000) and Svensson and Woodford (2003). Instrument rules go back much further. The catalyst for the most recent spate of papers on the subject was the papers of Henderson and McKibbin (1993) and Taylor (1993). Both strands of the literature are lengthy.

[2] Sometimes the loss function is an *ad hoc* one, said to represent the policy maker's preferences independent of the model to which the loss function is to be applied, as in Williams (2003) for example. In other cases, the loss function is a quadratic approximation to the representative agent's utility function in a dynamic stochastic general equilibrium (DSGE) model, as in Rotemberg and Woodford (1999).

[3] A linear model, or a non-linear model that has been linearized around its steady state. Recent contributions to the literature argue against linearization on the grounds that welfare comparisions of alternative policies on linearized models are often false; see, e.g., Kim and Kim (2003). See, however, Benigno and Woodford (2004) for a defense of linearization.

[4] In the case of what Svensson calls optimal targeting rules, the form of the rule is given by the structure of the model. For optimal instrument rules, the form of the rule is simplified in a form determined by the researcher on practical grounds, usually having to do with communications or tractability.

[5] The analysis of parameter uncertainty usually involves conditioning the optimal policy on the standard errors of estimated model coefficients. The seminal reference is Brainard (1968). More recent contributions include Sack (1999) and Soderstrom (2002). Data uncertainty obliges policy makers to take into explicit account pending data revisions. See Croushore and Stark (2001) and Rudebusch (2003).

forms by Hansen and Sargent (1995, 2004), Giannoni (2002), Onatski (2003) and Tetlow and von zur Muehlen (2001, 2004). The second examines uncertainty from a real-time perspective, an approach adopted in this paper.

The real-time analysis of policy is useful and important for several reasons. First, by examining the historical record, the researcher captures the real thing, not what some econometrician thinks, *ex post*, the real thing might have been. This is particularly important in the real-world setting where models have to be rolled out with due consideration to a work and data-release schedule. Second, real-time analysis can uncover issues that the econometrician might not envision, such as idiosyncratic shocks that would be omitted from a statistical model but are nonetheless important for policy.[6] Third, in the present context of addressing model uncertainty, real-time analysis narrows the range of would-be rival models under consideration to a plausible set.

This paper considers, measures and evaluates real-time model uncertainty in the United States. In particular, we study 30 vintages of the Board of Governors' mainline macroeconomic model, FRB/US, that were used extensively for forecasting and policy analysis at the Fed from the model's inception in July 1996 until November 2003. To do this, we exploit archives of the model code, coefficients, databases and stochastic shock sets for each vintage. The period of study was one of tumultuous change in the U.S. economy with a productivity boom, a stock market boom and bust, a recession, the Asia crisis, the Russian debt default, corporate governance scandals and an abrupt change in fiscal policy. There were also 23 changes in the intended federal funds rate, 7 increases and 16 decreases. We document the changes in the model properties that occurred during this period–a surprisingly large and consequential set, it turns out–and identify the economic events that contributed to these changes. We compute optimal Taylor-type rules for each vintage. We compare these *ex ante* optimal rules against alternative rules, in particular an *ex post* optimal rule, and the original Taylor (1993) specification. We draw conclusions about model uncertainty and its implications for policy design.

This exercise goes a number of steps beyond previous contributions to the literature. First, unlike the rival models literature, it does not involve models of artificial economies compared in a laboratory environment. Research in this area, while important and useful, is limited by the set of rival models and how they were selected. Second, we go beyond the literature on parameter uncertainty. That literature assumes that parameters are random but the model is fixed over time; misspecification is a matter of sampling error. We explicitly allow the models to change over time in response not just to the data but the economic issues of the day. [7] Lastly, and most

[6] The authors' favorite example of this is Y2K, where a "forecastable shock" to the optimal decision rules of firms was in play and was important, but had never occurred in the historical dataset and would never again. Other examples are legion.

[7] There have been a number of valuable contributions to the real-time analysis of monetary policy issues. Most are associated with data and forecasting. See, in particular, the work of

important, the analysis we provide derives from models that were actually used to facilitate Fed decision making. To the best of our knowledge, no one has ever done this before.

The rest of this paper proceeds as follows. The second section begins with a discussion of the FRB/US model(s) in generic terms, and the historical archives. The third section compares model properties by vintage, using four models per year, one for each release of National Income and Product Accounts data. To do this, we document changes in real-time "model multipliers" and compare them with their *ex post* counterparts. We also outline the economic events that coincided with changes in model properties. The succeeding section computes optimized Taylor-type rules and compares these to commonly accepted alternative policies in a stochastic environment. The fifth section examines the stochastic performance of candidate rules for two selected vintages, the February 1997 and November 2003 models. A sixth and final section sums up and concludes. Missing from this paper, but covered in a companion article, Tetlow (2004b), is an analysis of robust policies in the context of real-time model uncertainty.

2 Thirty vintages of the FRB/US model

2.1 A description of the FRB/US model

The FRB/US model came into production in July 1996 as a replacement for the venerable MIT-Penn-SSRC (MPS) model that had been in use at the Board of Governors for many years. Disenchantment with MPS had been growing through the 1980s, as the Board staff saw that the questions they were being asked by senior managers and Board members were changing in a way that the model was not designed to answer. Straightforward forecasting questions were giving way, more and more, to policy analysis questions that involved explicit consideration of alternative formulations of expectations. Among the questions of increasing interest were those that concerned the design of monetary policy rules, as opposed to discretionary, quarter-to-quarter settings of the funds rate. At the same time, the academic literature had been posing important questions about expectations formation and about the identification and interpretation of large-scale macroeconometric models. To address these challenges, the staff included within the FRB/US model a specific expectations block, and with it, a fundamental distinction between intrinsic model dynamics (dynamics that are immutable to policy) and expectational dynamics (which policy can affect). In most instances, the intrinsic dynamics of the model were designed around representative agents choosing optimal paths for decision variables facing polynomial adjustment

Croushore and Stark (2001) and a whole conference on the subject details of which can be found at http://www.phil.frb.org/econ/conf/rtdconfpapers.html An additional, deeper layer of real-time analysis considers revisions to unobservable state variables, such as potential output; on this see Orphanides *et al.* (2000) and Orphanides (2001).

costs. The notion of polynomial adjustment costs, a straightforward generalization of the well-known quadratic adjustment costs, allowed, for example, the flow of investment to be costly to adjust, and not just the capital stock. This idea, controversial at the time, has recently been adopted in the broader academic community.[8]

The structure of macroeconomic models at the Fed have always responded to economic events and the different questions that those events evoke, even before FRB/US. Brayton, Levin, Tryon and Williams (1997) note, for example, how the presence of financial market regulations meant that for years a substantial portion of the MPS model dealt specifically with mortgage credit and financial markets more broadly. The repeal of Regulation Q induced the elimination of much of that detailed model code. Earlier, the oil price shocks of the 1970s and the collapse of Bretton Woods gave the model a more international flavor than it had previously. We shall see that this responsiveness of models to economic conditions and questions continued with the FRB/US model in the 1990s.

From the outset, FRB/US has been a significantly smaller model than was MPS, but it is still fairly large. At inception, it contained some 300 equations and identities of which perhaps 50 were behavioral. About half of the behavioral equations in the first vintage of the model were modeled using formal specifications of optimizing behavior containing explicit estimates of forward expectations and adjustment costs.[9] As already noted, among the identities are equations governing expectations formation.

Two versions of expectations formation were envisioned: VAR-based expectations and perfect foresight. The concept of perfect foresight is well understood, but VAR-based expectations probably requires some explanation. The parable told when the model is simulated in VAR model, is very much like the Phelps-Lucas "island paradigm": the models agents live on different islands where they have access to a limited set of core macroeconomic variables, knowledge they share with everyone in the economy. The core macroeconomic variables are the output gap, the inflation rate and the federal funds rate, as well as beliefs on the long-run target rate of inflation and the what the equilibrium real rate of interest will be in the long run. In addition they have information that is germane to their island, or sector. Consumers, for example, augment their core VAR model with information about potential output growth and the ratio of household income to GDP. Perturbations to the model under VAR-based expectations are treated as generic shocks, the implications of which play out over time in a way determined by the broad dynamics of the model.

By definition, under perfect-foresight expectations, the information set is broadened to include all the states in the model with all the cross-equation restrictions implied by the model. The duration and implications of perturbations under perfect

[8] Christiano, Eichenbaum and Evans (2004), for example, allow the flow of investment to be costly to adjust which is the same thing as having higher-order adjustment costs for the stock of capital.

[9] Polynomial adjustment costs in price and volume decision rules. In financial markets, intrinsic adjustment costs were assumed to be zero.

foresight are normally treated as known. Those who use the model have come to think of the perfect-foresight version of the model as being what the VAR-based expectations version would converge upon if a constant monetary policy rule were in place for a lengthy period of time and agents on their islands were able to gradually learn off-island information. Typically, the VAR-based expectations version of the model is used for forecasting, and for policy analysis when the staff believes the experiment in question does not deviate too much from what has been typical in the past so that the average historical experience captured in the VAR can be thought of as representative of the response is likely to be under the experiment. The perfect-foresight version is used for problems in which agents are likely to have the information and motivation to formulate a detailed understanding of events.[10] In the end, with either version, agents' decision rules (except with asset prices) usually end up looking like hybrid New Keynesian model equations.

There is not the space here for a complete description of the model, a problem that is exacerbated by the fact that the model is a moving target. Readers interested in detailed descriptions of the model are invited to consult papers on the subject, including Brayton and Tinsley (1996), Brayton, Levin, Tryon and Williams (1997), and Reifschneider, Tetlow and Williams (1999). Our discussion here is more stylized. Ignoring asset pricing equations, a generic model equation would look something like:

$$\Delta x = \alpha(L)\Delta x + E_t\beta(F)\Delta x^* + c(x_{t-1} - x^*_{t-1}) + u_t \tag{1}$$

where $\alpha(L)$ is a polynomial in the lag operator, i.e., $\alpha(L)z = a_0 + a_1 z_{t-1} + a_2 z_{t-2} + ...$ and $\beta(F)$ is a polynomial in the *lead* operator. The term Δx^* is the expected changes in target levels of the generic decision variable, x, $c(.)$ is an error-correction term, and u is a residual. In general, the theory behind the model will involve cross-parameter restrictions on $\alpha(L)$,$\beta(F)$ and c. The point to be taken from equation (1) is that decisions today for the variable, x, will depend in part on past values and future values, with an eye on bringing x toward its desired value, x^*, over time.

The main objectives guiding the development of the model were that it be useful for both forecasting and policy analysis; that expectations be explicit and important equations representing the decision rules of optimizing agents; that the model be estimated and have satisfactory statistical properties; and that the full-model simulation properties match the "established rules of thumb regarding economic relationships under appropriate circumstances."[11]

Our concern in this paper is the monetary transmission mechanism and how uncertainty about it can affect policy. The key features influencing this in the FRB/US model are the effects of changes in the funds rate on asset prices and from there to expenditures. Philosophically, the model has not changed much in this area. All

[10] Examples of where foresight is regarded as critical include certain kinds of fiscal policy interventions since they involve legislative commitments to future actions that are costly to undo and for which it pays for agents to make the effort to learn the implications of the legislation.

[11] Brayton and Tinsley (1996), p. 2.

vintages of the model have had expectations of future economic conditions in general, and the federal funds rate in particular, affecting long-term interest rates and inflation. From this, real interest rates are determined and this in turn affects stock prices and exchange rates, and from there, real expenditures. Similarly, the model has always had a wage-price block, with the same basic features: sticky wages and prices, expected future excess demand in the goods and labor markets influencing price and wage setting, and a channel through which productivity affects real and nominal wages. That said, as we shall see, there have been substantial changes over time in both (what we may call) the interest elasticity of aggregate demand and the effect of excess demand on inflation.

Over the years, equations have come and gone in reflection of the needs, and data, of the day. The model began with an automotive sector but this block was later dropped. Business fixed investment was originally disaggregated into just non-residential structures and producers' durable equipment, but the latter is now disaggregated into high-tech equipment and "other". The key consumer decision rules and wage-price block have undergone almost continuous change over the period. On the other hand, the model has always had a decision rule for consumer non-durables and services, one for consumer durables expenditures, and one for housing. There has always been a trade block, with aggregate exports and non-oil and oil imports, and equations for foreign variables. The model has always had a three-factor, constant-returns-to-scale Cobb-Douglas production function with capital, labor hours and energy as factor inputs.

2.2 The archive

Since its inception in July 1996, the FRB/US model code, the equation coefficients, the baseline forecast database, and the list of stochastic shocks with which the model would be stochastically simulated, have all been stored for each of the eight forecasts the Board staff conducts every year. In principle, this means that 60 model vintages are available. In order to match the models to the quarterly frequency of the data, we elected to use four archives per year, the ones immediately following National Income and Product Accounts preliminary releases. This ensures that we are considering models for which data releases may have elicited a significant change. There are 30 such vintages with which we can work.[12]

In what follows, we experiment with each vintage of model, comparing their properties in selected experiments. Consistent with the real-time philosophy of this endeavor, the experiments we choose are typical of those used to assess models by policy

[12]The archives are listed by the precise date of the FOMC meeting in which the forecasts were discussed. For our purposes, we do not need to be so precise so we shall describe them by month and year. Thus, the 30 vintages we use are, in 1996: July and November; in 1997: February, May, July, and November; in 1998 through 2000: February, May, August and November; and in 2001 through 2003: January, May, August and November.

institutions in general and the Federal Reserve Board in particular. They fall into two broad classes. One set of experiments, *model multipliers*, attempt to isolate the behavior of particular parts of the model. A multiplier is the response to an exogenous shock of a key endogenous variable after a fixed period of time. An example is the response of the unemployment rate after eight quarters to a persistent increase in the federal funds rate. We shall examine several such multipliers. The other set of experiments judge the stochastic performance of the model and are designed to capture the full-model properties under fairly general conditions. So, for example, we will be computing by stochastic simulation the optimal coefficients of a Taylor rule, conditional on a model vintage, a baseline database, and a set of stochastic shocks.[13] We will then compare these optimal rules with other alternative rules and indeed other alternative worlds defined by the set contained in our model vintages. To the extent that any of these givens turn out not to be representative of the *ex post* experience, the *ex ante* optimal Taylor rule will turn out not to be optimal *ex post*.

The archives document model changes and provide a unique record of model uncertainty. As we shall see, the answers to questions a policy maker might ask are different depending on the vintage of the model. The seemingly generic issue of the output cost of bringing down inflation, for example, can be subdivided into several more precise questions, including: (i) what would the model say is the output cost of bringing down inflation today?; (ii) what would the model say would have been the output cost of bringing down inflation in February 1997?; and (iii) what would the model have said in February 1997 was the output cost of disinflation? These questions introduce a time dependency to the issue that rarely appears in other contexts. The answers to these and other related questions depend on the model vintage. Here, however, the model vintage means more than just the model alone. Depending on the question, the answer can depend on the baseline; that is, on the initial conditions from which a given experiment is carried out. It can also depend on the way an experiment is carried out, and in particular on the policy rule that is in force. And since models are evaluated in terms of their stochastic performance, it can depend on the stochastic shocks to which the model is subjected to judge the appropriate policy and to assess performance. So in the most general case, model uncertainty in our context comes from four interrelated sources: model, policy rule, baseline and stochastic shocks.

How much model variability can there be over a period of just eight years? The answer is a surprisingly large amount. But to provide a specific answer, let us begin with the data. It is ultimately the data that underlie changes in the model, changes in the stochastic shocks, and changes in the policy rules that react to those shocks and control the model. In the spirit of Orphanides (2001), let us begin by examining

[13] Each vintage has a list of variables that are shocked using bootstrap methods for stochastic simulations. The list of shocks is a subset of the model's complete set of residuals since other residuals are treated not as shocks but rather as measurement error. The precise nature of the shocks will vary according to data construction and the period over which the shocks are drawn.

measures of historical data by vintage. Figure 1 shows the four-quarter growth rate of the GDP price index, for selected vintages. (Note we show only real-time historical data because of rules forbidding the publication of forecast data more recent than in the last five years.) The inflation rate moves around some, but the various vintages for the most part are highly correlated. That said, our reading of the literature is that data uncertainty, narrowly defined to include revisions of published data series, is not a first-order source of problems for monetary policy design; see, e.g., Croushore and Stark (2001). As argued by Orphanides *et al.* (2000) and Orphanides (2001), unobservable variables like potential output are, or at least may be, another story.

Figure 2 shows the more empirically important case of model measures of growth in potential non-farm business output. Unlike the case of inflation, potential output growth is a latent variable the definition of which depends on model concepts. What this means is the historical measures of potential are themselves a part of the model–so we should expect significant revisions.[14] Even so, the magnitudes of the revisions shown in Figure 2 are truly remarkable. The July 1996 vintage shows potential output growth of about 2 percent. For the next several years, succeeding vintages show both higher potential output growth rates and more responsiveness to economic developments. By January 2001, growth in potential was estimated at over 5 percent for some dates, before subsequent changes resulted in a path that was lower and more variable. Why might this be? Table 1 reminds us about how extraordinary the late 1990s were. The table shows selected FRB/US model forecasts for the four-quarter growth in GDP over the period for which public availability of the data are not restricted.[15] The table shows the substantial underprediction of GDP growth over the period, an experience that was common among forecasters. As the underpredictions persisted, the staff came, more and more, to view the shocks underlying the rapid growth (with steady to falling inflation) as representing persistent shocks to the growth rate of productivity. Consequently, the staff began adding model code to allow the supply side of the model to respond to output surprises by projecting forward revised profiles for productivity growth. The result is the *ex post* correlation of potential output growth and the business cycle shown in Figure 2.

[14] Defined in this way, data uncertainty does *not* include uncertainty in the measurement of latent variables, like potential output. The important conceptual distinction between the two is that eventually one knows what the final data series is–what "the truth" is–when dealing with data uncertainty. One never knows, even long after the fact, what the true values of latent variables are. Latent variables are more akin to parameter uncertainty than data uncertainty.

[15] A record such as the one in the table was not unusual during this period; the Survey of Professional Forecasters similarly underpredicted output growth.

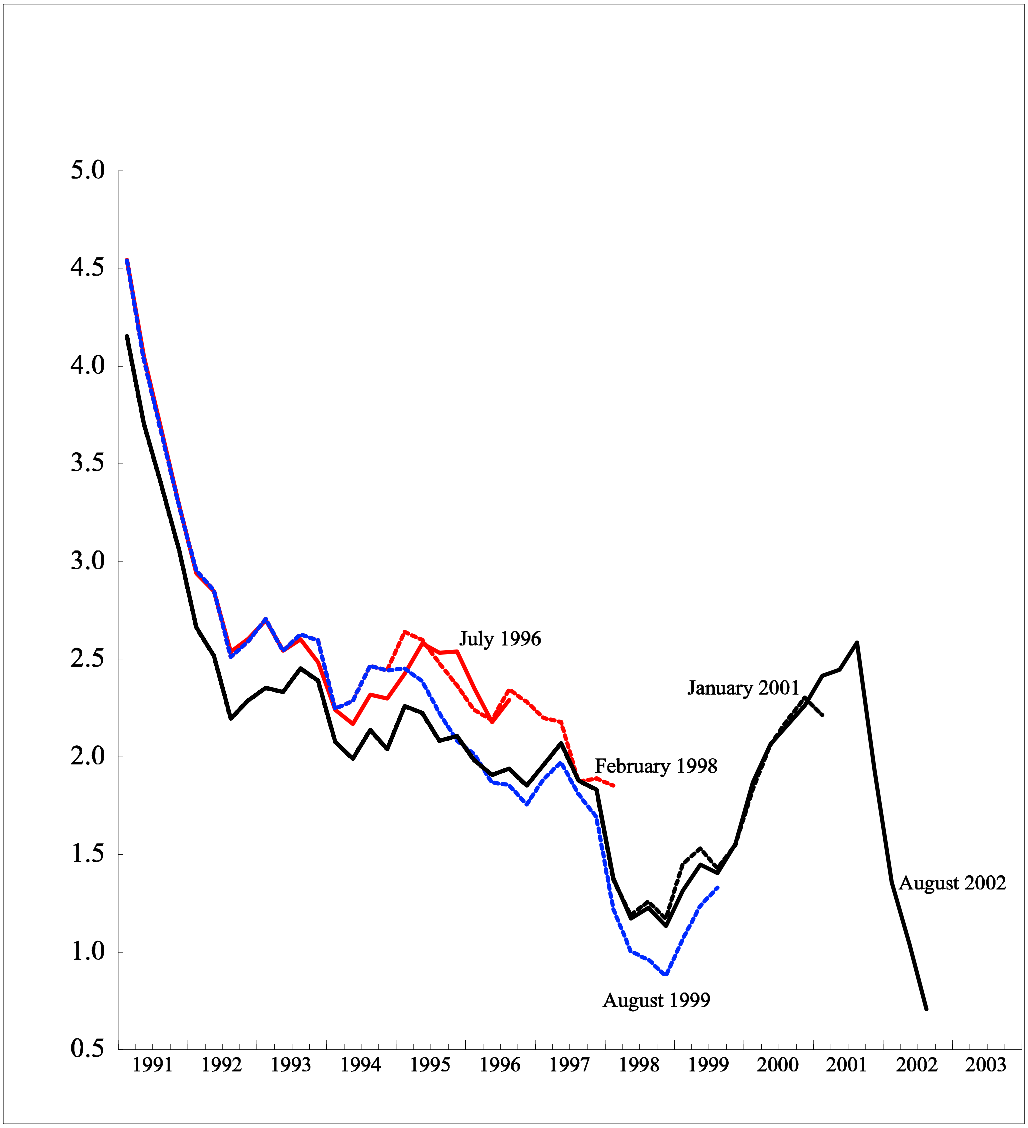


Figure 1: Real-time 4-quarter GDP price inflation

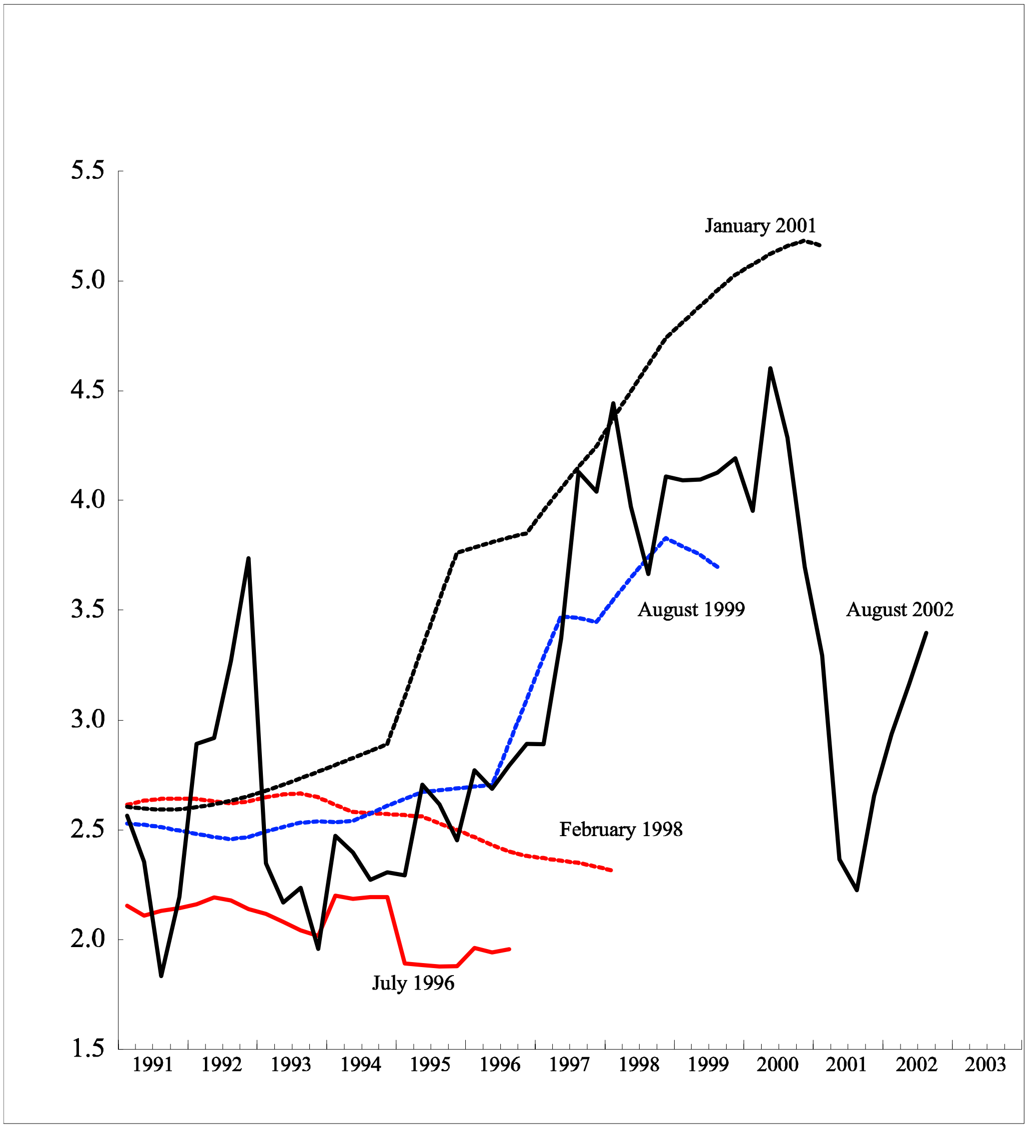


Figure 2: Real-time 4-quarter non-farm potential output growth

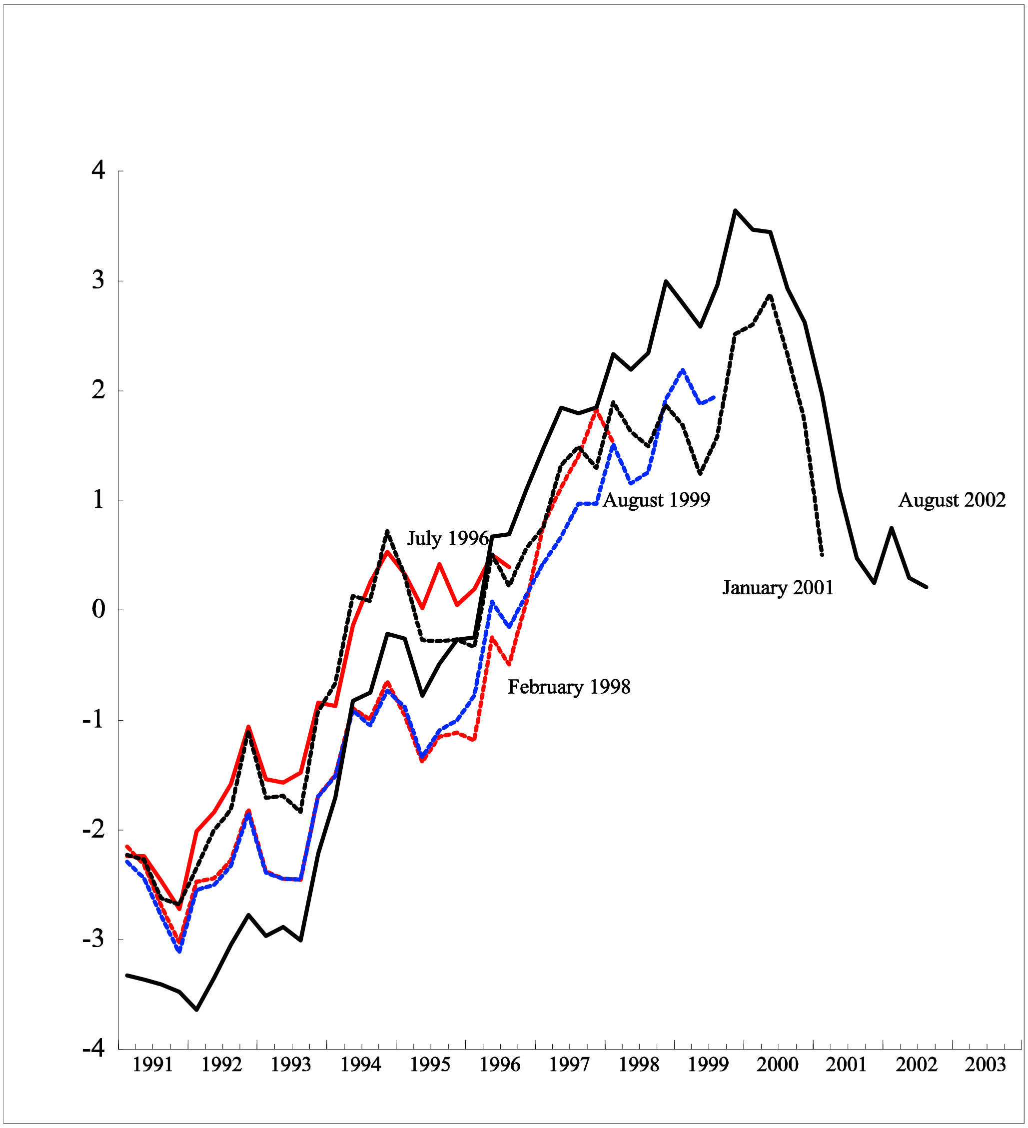


Figure 3: Real-time GDP output gaps

Table 1			
Q4 growth in GDP: selected FRB/US model forecasts			
forecast date	forecast	final data	actual less final
July 1996	2.2	4.8	2.6
July 1997	2.0	3.7	1.7
Aug.1998	3.0	4.4	1.4
Aug.1999	3.2	3.6	0.4
4Q GDP growth forecasts from the third-quarter vintage of the year shown compared against final data.			

The most recent historical measures shown in the figure are for the August 2002 vintage, where the path for non-farm potential output growth differs in several ways from the others. The first way in which it differs is that it is the only series that is less optimistic than earlier ones In part, this reflects the onset of the 2001 recession. The second way the series differs is in its volatility over time. This is a manifestation of the ongoing evolution of the model in response to economic conditions. In its early vintages, the modeling of potential output in FRB/US was traditional for large-scale econometric models, in that trend labor productivity and trend labor input, were based on exogenous split time trends. Beginning with the May 2001 model, a production function approach was adopted which allowed capital services to play a direct role in the evolution of potential. This shows up in more volatile measures of potential, among other things.

Despite the volatility of potential output growth, the resulting output gaps, shown in Figure 3, show considerable covariation, albeit with non-trivial variation over time.

Figures 4 and 5 which are really one figure spread over two pages, provide a helicopter tour of the model's changes over time and juxtaposes events that might have elicited those changes. The chart across the top shows two things: the total number of equation changes by vintage (the red bars, measured off the left-hand scale), and the total number of model equations, including identities (the blue line and the right-hand scale). Three facts immediately arise from the picture. First, there have been flurries of numerous changes in the model at times. Second, the number of changes has tended to decrease over time.[16] And third, the number of equations has increased, particularly in the period from 2000 to 2002. The fact that many model changes were undertaken early in the model's history but without adding to the size of the model while fewer changes were adopted later on that nonetheless added to the model's size suggests that early period was one of model refinement while the latter period was one of reform. Indeed, during the period from about 1998 to 2002, the

[16] A "model change" is the non-trivial addition, deletion or change in specification of a significant model equation from the vintage immediately preceding. Re-estimation of a given equation does not count as a model change. Rewriting an equation in a mathematically equivalent way is not counted. In a fully articulated model with a large number of identities, changes in structural equations can oblige corresponding changes in a large number of associated identities. As a result, the count of model changes mounts rapidly.

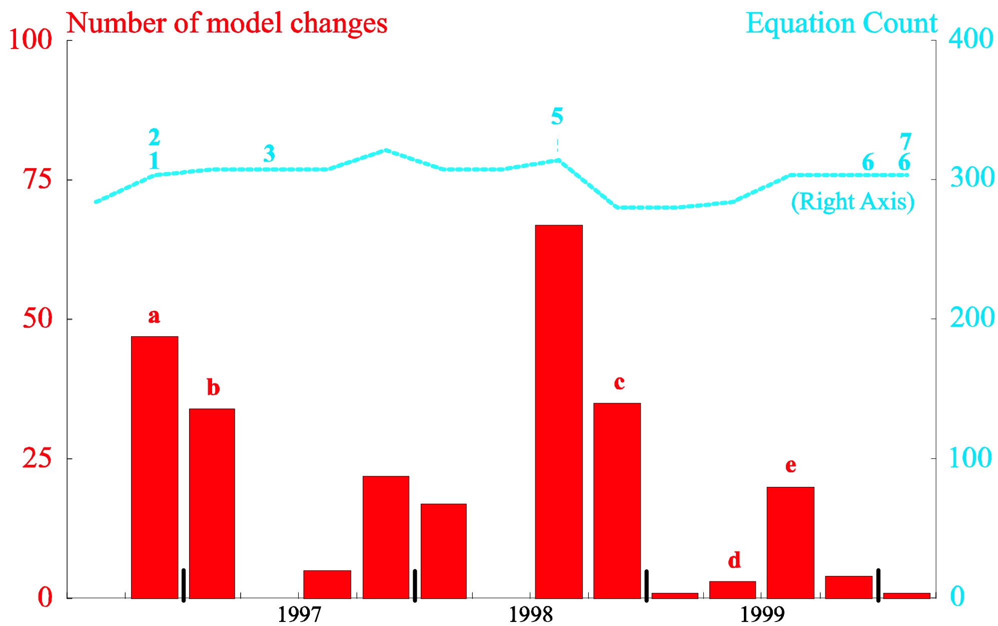
range of questions that the model was expected to address increased, and the staff's view of the economy became more complicated. The bottom part (of both pages) of the figure identifies some factors that may have played a role in the model changes, as well as providing some reminders of the events of the time. The left-hand column of each page is devoted to significant events over the period; each entry is marked by a number, with a corresponding entry appearing in the appropriate place, in the same color, in the chart. The right-hand column identifies selected model changes; these are marked by a letter, with a corresponding mark on the chart.

Looking at the table, we can see from the outset that economic forces of the day were influencing the model's specification. The stock market was already booming in July 1996, when the model was brought into service. By the end of the year, the model's stock market equation and the consumption and housing equations that stock market wealth affect had been changed. The most significant changes came, however, as the lasting implications of the productivity boom became prominent. In December 1999, as a part of the comprehensive revisions to the National Income and Product Accounts, software was added to the measurement of the capital stock.[17] Investment expenditures–particularly expenditures on information technology–boomed over the same period as did stock market valuations. By late 1999, it became clear that machinery and equipment expenditures would have to be disaggregated into high-tech and "other". The boom also engendered other questions: what is the effect of an acceleration in productivity on the equilibrium real interest rate and on the savings rate? What are the implications of a permanent shift in the relative price of computer equipment? These and other questions resulted in a reformulation of the model's supply side. The unifying theme of the questions of the time was a longer run or lower frequency orientation than had previously been the case. The introduction of chain-weighted data in late 1996 made modeling these low-frequency trends feasible in a way that had not been the case before.[18] The point is that changes to the model were not always a reflection of the model underperforming at the tasks it was originally built to do; in many instances, it was an outcome of an expansion of the tasks to which the model was assigned.

To summarize this section, the FRB/US model archives show considerable change in equations and the data by vintage. The next section examines the extent to which these differences manifest themselves in different model properties. The following section then examines how these differences, together with their associated stochastic shock sets, imply different optimal monetary policy rules.

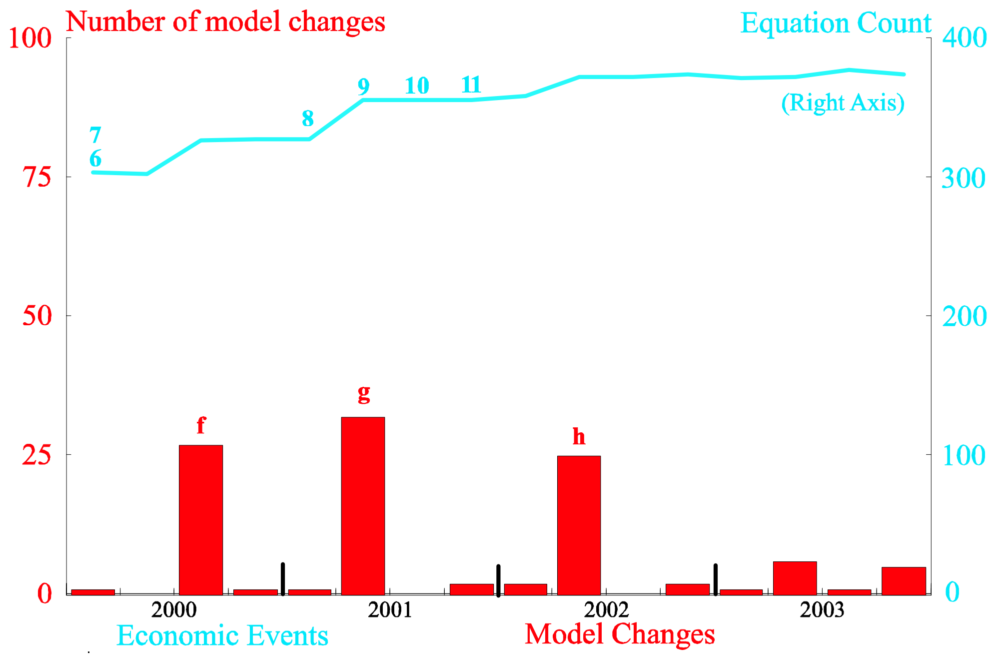
[17]Prior to that time, expenditures on software were regarded as an intermediate input; they had no direct effect on GDP.

[18]In the absence of chain-weighting, trends in relative prices, like the relative price of high-tech capital goods, could not be modeled well. The inability to account for weight shifts in expenditure bundles, which was merely a nuisance over short horizons, was a substantial barrier for the analysis of longer-term phenomena.



	Economic Events	Model Changes
1996:Q4	1 - "Irrational exuberance" speech 2 - BEA unveils chain-weighted data	**a** - Consumption and housing block refined - Foreign sector revised - Change in stock market equation
1997:Q1		**b** - Trends revised in the labor sector - Bond rate term allowed to time-vary
1997:Q2	3- Asia crisis begins	
1998:Q3	4- Russian debt default 5 - Fall of LTCM	
1998:Q4		**c** - Housing equation revised - Foreign sector revised as a result of a shift from G10 to G29 aggregates - Main price equation revised as a result
1999:Q2		**d** - Stock market equation revised from constant-elasticity to linear form
1999:Q3		**e** - Chain aggregation introduced to model
1999:Q4	6 - NIPA comprehensive Revisions I: software added to business fixed investment	
2000:Q1	6 - NIPA revisions II: software added to capital stock 7 - Stock market peak	

Figure 4: Model changes by vintage, 1997-1999



2000:Q1	6 - NIPA revisions II: software added to capital stock 7 - Stock market peak	
2000:Q3		**f** - Major changes in response to NIPA revisions - new investment block - wage-price sector revised - relative prices term added to consumer durables equation - equipment, software, and inventories equation revised to include an effect of the relative price of computers - stock market equation adjusted
2001:Q1	8 - Start of recession	
2001:Q2	9 - First Bush budget passes	**g** - Equipment and software disaggregated into high tech and "other" components - New supply side links between investment and labor - wage and price equations estimated simultaneously using FIML
2001:Q3	10 - September 11th	
2001:Q4	11 - Enron collapses	
2002:Q2		**h** - Stochastic trends added to labor sector

Figure 5: Model changes by vintage, 2000-2003

3 Model multipliers in real time and *ex post*

In this subsection, we consider the variation in real time of selected model multipliers. In most instances, we are interested in the response after 8 quarters of unemployment to a given shock, (although our first experiment is an exception to this rule). We choose unemployment as our response variable because it is one of the two key real variables that the Fed has concerned itself with over the years, and the other, the output gap, has changed definitions over time. We would like to avoid differences in multipliers that arise solely from changes in definitions. The horizon of eight quarters is a typical one for exercises such as this as conducted at the Fed and other policy institutions. Except where otherwise noted, we hold the nominal federal funds rate at baseline for each of these experiments.

It is easiest to show the results graphically. But before turning to specific results, it is useful to outline how these figures are constructed and how they should be interpreted. In all cases, we show two lines. The black solid line is the real-time multiplier by vintage. Each point on the line represents the outcome of the same experiment, conducted on the model vintage of that date, using the baseline database at that point in history. So at each point shown by the black line, the model, its coefficients and the baseline all differ. The red dashed line shows what we call the *ex post* multiplier. The *ex post* multiplier is computed using the most recent model vintage for each date; the only thing that changes for each point on the dashed red line is the initial conditions under which the experiment is conducted. Differences over time in the red line reveal the extent to which the model is nonlinear, because the multipliers for linear models are independent of initial conditions.

Now let us look at Figure 6, which shows the 5-year employment sacrifice ratio; that is, the cost in terms of cumulative annualized forgone employment, that a one-percentage-point reduction in the inflation rate would cost after five years.[19] Let us focus on the red dashed line first. It shows that for the November 2003 model, the sacrifice ratio is essentially constant over time. So if the staff were asked to assess the sacrifice ratio, or what the sacrifice ratio would have been in, say, February 1997, the answer based on the November 2003 model would be the same: about 4-1/4, meaning that it would take that many percentage-point-years of unemployment to bring down inflation by one percent. Now, however, look at the black solid line. Since each point on the line represents a different model, and the last point on the far right of the line is the November 2003 model, the red dashed line and the black solid line must meet at the right-hand side in this and all other figures in this section. But notice how much the real-time sacrifice ratio has changed over the 8-year period of study. Had the model builders been asked in February 1997 what the sacrifice ratio

[19] More precisely, the experiment is conducted by simulation, setting the target rate of inflation in a Taylor rule to one percentage point below its baseline level, and setting the feedback coefficient on the output gap to zero. The sacrifice ratio is cumulative annualized change in the unemployment rate, undiscounted, relative to baseline, divided by the change in PCE inflation after 5 years.

was, the answer based on the February 1997 model would have been about 2-1/4, or approximately half the November 2003 answer. The black line undulates a bit, but cutting through the wiggles, there is a general upward creep over time, and a fairly discrete jump in the sacrifice ratio in late 2001.

The climb in the model sacrifice ratio is striking, particularly as it was incurred over such a short period of time. One might be forgiven for thinking that this phenomenon is idiosyncratic to the model under study. On this, two facts should be noted. First, even if it were idiosyncratic such a reaction misses the point. The point here is that these are the models that were used by the Fed staff and they were constructed with all due diligence to address the sort of questions asked here. Second, work in progress to be documented later shows that this result is not a fluke: The same phenomenon occurs to varying degrees in simple single-equation Phillips curves of various specifications using both real-time and *ex post* data. [20]

Figure 7 shows the funds-rate multiplier; that is, the increase in the unemployment rate after eight quarters in response to a persistent 100-basis-point increase in the funds rate. This time, the red dashed line shows important time variation: the *ex post* funds rate multiplier varies with initial conditions, it is highest at a bit over 1 percentage point in late 2000, and lowest at the end of the period. The nonlinearity stems entirely from the linearity (as opposed to log linearity) of the model's stock market equation which makes the interest elasticity of aggregate demand an increasing function of the share of stock market wealth to total wealth. The mechanism is that an increase in the funds rate raises long-term bond rates, which in turn bring about a drop in stock market valuation operating through the arbitrage relationship between expected risk-adjusted bond and equity returns.

The real-time multiplier, shown by the solid black line is harder to characterize. Two observations stand out. The first is the sheer volatility of the multiplier. In a large-scale model such as the FRB/US model, where the transmission of monetary policy operates through a number of channels, time variation in the interest elasticity of aggregate demand depends on a large variety of parameters. Second, the real-time multiplier is almost always lower than the *ex post* multiplier. The gap between the two is particularly marked in 2000, when the business cycle reached a peak, as did stock prices. At the time, concerns about possible stock market bubbles were rampant. One aspect of the debate between proponents and detractors of the active approach to stock market bubbles centers around the extent to which the proper policy approach to bubbles is feasible in a world of model uncertainty. Our results suggest that this concern is a real one.[21] In fact, there were three increases in the federal funds rate

[20] Cogley and Sargent (2004) use Bayesian techniques to estimate two Phillips curves and an aggregate supply curve simultaneously asking why the Fed did not choose an inflation stabilizing policy before the Volcker disinflation. They too find time variation in the (reduced-form) output cost of disinflation. Roberts (2004) shows how greater discipline in monetary policy may have contributed to the reduction in economic volatility in the period since the Volcker disinflation.

[21] The "active approach" to the presence of stock market bubbles argues that monetary policy should specifically respond to bubbles. See, e.g., Cecchetti *et al.* (2000). The passive approach

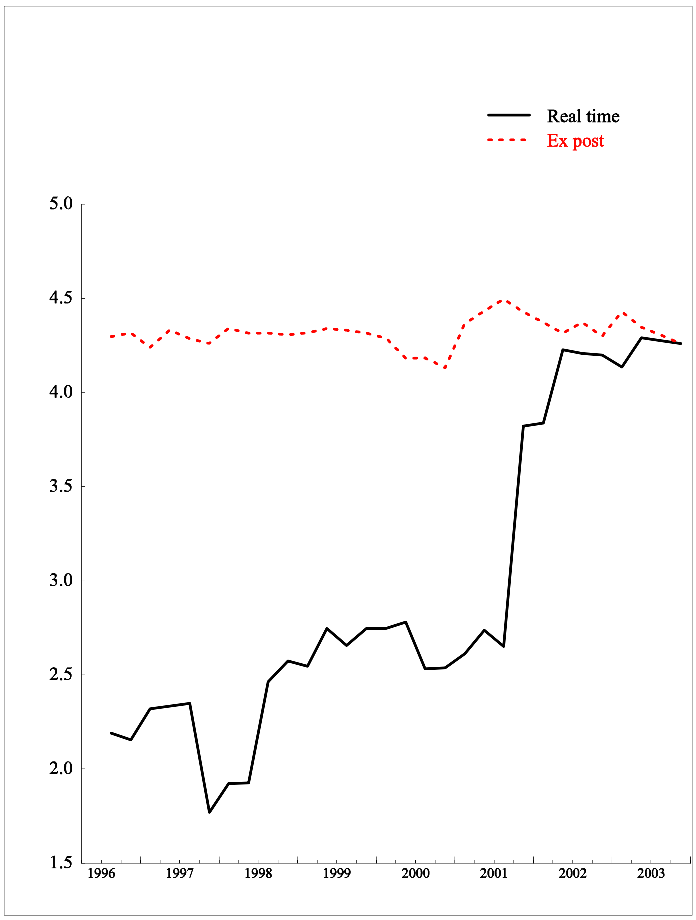


Figure 6: Sacrifice ratio by model vintage

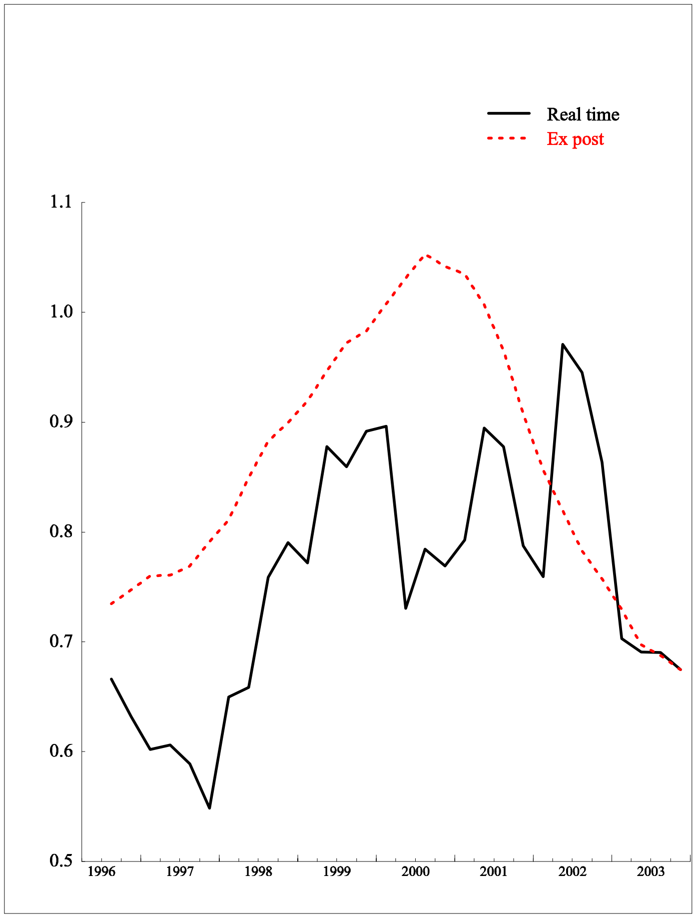


Figure 7: Funds rate multiplier by model vintage

during 2000, totalling 100 basis points.[22] Once again, a counterfactual analysis of the strength of monetary policy historically based on the *ex post* model could render misleading answers.

Figure 8 shows the government expenditure multiplier–the effect on the unemployment rate of a persistent increase in government spending of 1 percent of GDP. Noting that the sign on this multiplier is negative, one aspect of this figure is the same as the previous one: the real-time multiplier is nearly always smaller (in absolute terms) than *ex post* multiplier. If we take the *ex post* multiplier as correct, this says that a policy maker relying on the real-time estimates through recent history would have routinely underestimated the extent to which perturbations in fiscal

argues that bubbles should affect monetary policy only insofar as they affect the forecast for inflation and possibly output. They should not be a special object of policy. See, Bernanke *et al.* (1999, 2001), and Tetlow (2004a).

[22] The intended federal funds rate was raised 25 basis points on February 2, 2000, to 5-3/4 percent; by a further 25 basis points on March 21, and by 50 basis points on May 16, to 6-1/2 percent.

policy would oblige an offsetting monetary policy response. Given that the period of study involved a substantial change in the stance of fiscal policy, this is an important observation. A second aspect of the figure is the near-term reduction in the *ex post* multiplier, from about -0.9 in the 1990s, to about -0.75 in this decade.

Lastly, Figure 9 shows the effects on unemployment of a persistent change in the trend growth rate of productivity. In this instance, the forementioned non-linearity influences the *ex post* multiplier. The shock has its largest effects on unemployment when the stock market is a large proportion of household wealth. Somewhat surprisingly, given the flurry of respecifications that the productivity boom elicited in the late 1990s, the real-time multiplier differs from the *ex post* multiplier in only small ways. Evidently, it is the productivity boom's effect operating through the stock market that has been the key to model properties regardless of model vintage.[23]

To summarize this section, real-time multipliers show substantial variation over time, and differ considerably from what one would say e*x post* the multipliers would be. Moreover, the discrepancies between the two multiplier concepts have often been large at critical junctures in recent economic history. It follows that real-time model uncertainty is an important problem for policy makers. The next section quantifies this point by characterizing optimal policy, and its time variation, conditional on these model vintages.

4 Monetary policy in real time

4.1 Optimized Taylor rules

One way to quantify the importance of model uncertainty for monetary policy is to examine how policy advice would differ depending on the model. A popular device for providing policy advice is with the prescribed paths for interest rates from simple monetary policy rules, like the rule proposed by Taylor (1993) and Henderson and McKibbin (1993). A straightforward way to do this is to examine compute optimized Taylor (1993) rules. Many central banks use simple rules of one sort or another in the assessment of monetary policy and for formulating policy advice. Because they react to only those variables that would be key in a wide set of models, simple rules are said to be robust to misspecification. Thus, optimized Taylor rules would appear to be an ideal vehicle for study.

Formally, a Taylor rule is optimized by choosing the parameters of the rule, $\Phi = \{\alpha_Y, \alpha_\Pi\}$ to minimize a loss function subject to a given model, $x = f(\cdot)$, and a given set of stochastic shocks, Σ:

[23] Figures 6 to 9 capture the salient points about real-time model uncertainty fairly well. That said, we examined a number of other multipliers. Graphs of these are available from the authors upon request.

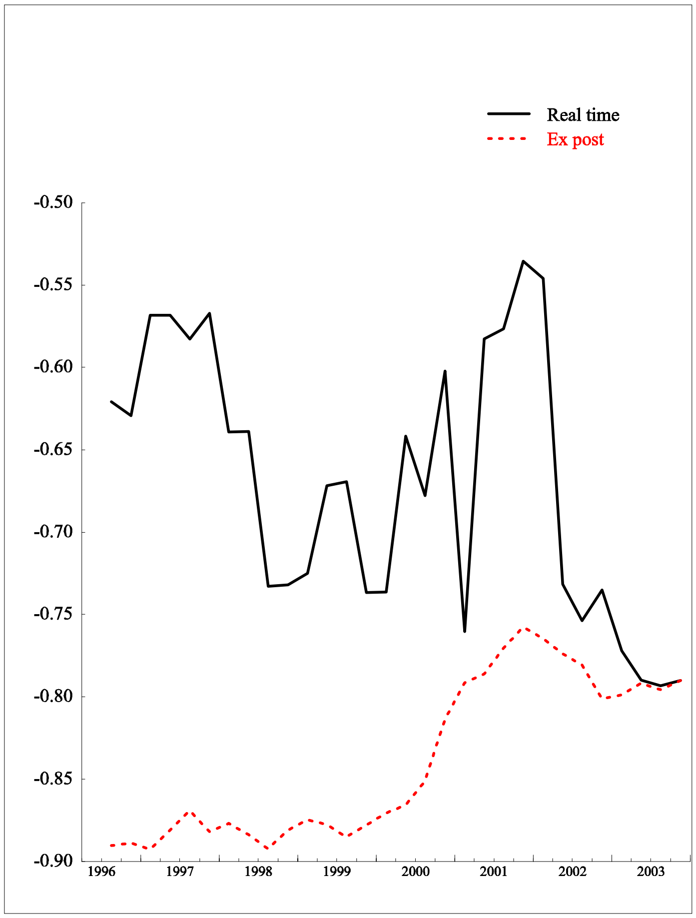


Figure 8: Government expenditure multiplier by vintage

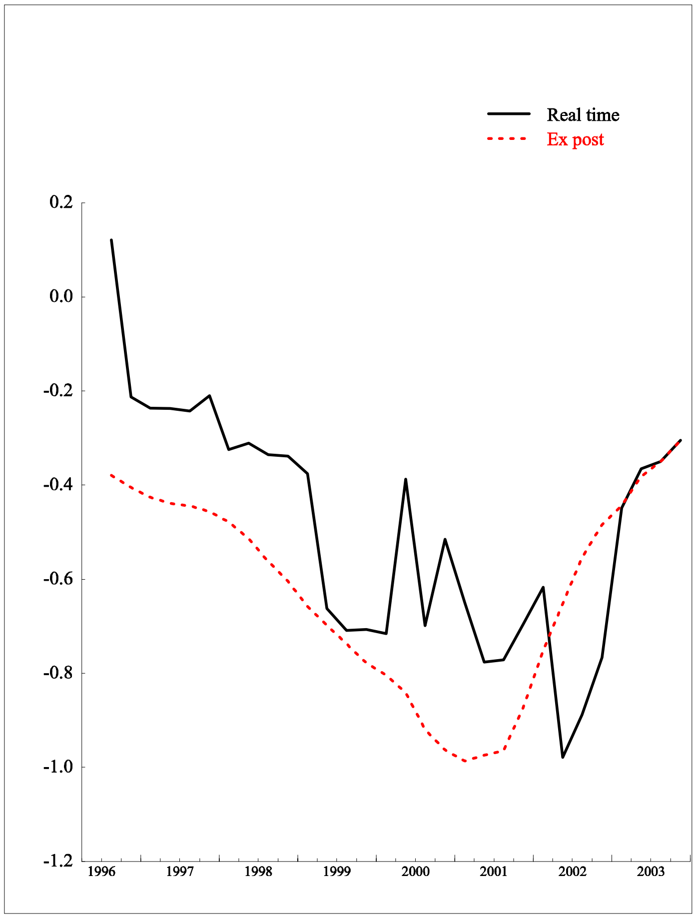


Figure 9: Productivity growth multiplier by vintage

$$\underset{\langle\Phi\rangle}{MIN} \sum_{i=0}^{T} \beta^i \left[\left(\pi_{t+i} - \pi^*_{t+i}\right)^2 + \lambda_Y \left(y_{t+i} - y^*_{t+i}\right)^2 + \lambda_{\Delta R} (\Delta r_{t+i})^2 \right] \tag{2}$$

subject to:

$$x_t = f(x_t,...x_{t-j}, z_t,...z_{t-k}, r_t,...r_{t-m}) + u_t \qquad j,k,m > 0 \tag{3}$$

and

$$r = rr^*_t + \widetilde{\pi} + \alpha_Y (y_t - y^*_t) + \alpha_\Pi (\widetilde{\pi}_t - \pi^*_t) \tag{4}$$

and

$$\Sigma_u = u'u \tag{5}$$

where x is a vector of endogenous variables, and z a vector of exogenous variables, both in logs, except for those variables measured in rates, π is the inflation rate, $\widetilde{\pi} = \Sigma_{i=0}^{3} \pi_{t-i}/4$ is the four-quarter moving average of inflation, π^* is the target rate of inflation, y is (the log of) output; y^* is potential output, and r is the federal funds rate. Trivially, it is true that: $\pi, \pi^*, y, y^*, \Delta r \in x$.[24],[25] In principle, the loss function, (2), could have been derived as the quadratic approximation to the true social welfare function for the FRB/US model. However, it is technically infeasible for a model the size of FRB/US. That said, with the possible exception of the term penalizing the change in the federal funds rate, the arguments to 2 are standard. The penalty on the change in the funds rate may be thought of as representing either a hedge against model uncertainty by the Fed, which may wish to reduce the likelihood of the fed funds rate entering ranges beyond those for which the model was estimated, or as a pure preference of the Committee. Whatever the reason for its presence, the literature confirms that some penalty is needed to explain the historical persistence of monetary policy; see, e.g., Sack and Wieland (2000) and Rudebusch (2001).

Solving a problem like this can be done easily for linear models. FRB/US is, however, a non-linear model. We therefore compute the optimized rule by stochastic simulation. Specifically, each vintage of the model is subjected to bootstrapped shocks from its stochastic shock archive over each model's historical database. Historical shocks from the estimation period of the key behavioral equations are drawn.[26]

[24] The fact that the policy rule depends on the variance-covariance matrix of stochastic shocks means that the rule is not certainty equivalent. This is the case for two reasons. One is the non-linearity of the model. The other is the fact that the rule is a simple one: it does not include all the states of the model.

[25] The equilibrium real interest rate, rr^*, is an endogenous variable in the model. In particular, $rr^*_t = (1-\gamma) rr^*_{t-1} + \gamma (rn_t - \pi_t)$ where r is the federal funds rate, and γ=0.05. As a robustness check, we experimented with allowing an intercept in the optimized rules in addition to rr^* and always found that this term was virtually zero for every model vintage.

[26] The number of shocks used for stochastic simulations has varied with the vintage, and generally has grown. For the first vintage, 43 shocks were used, while for the November 2003 vintage, 75 were used.

400 draws of 80 periods each are used in simulation to evaluate candidate parameterizations, with a simplex method used to determine the search direction.

This is obviously a very computer intensive exercise and so we are limited in the range of preferences we can investigate. Accordingly, we discuss only for one set of preferences: equal weights on output, inflation and the change in the federal funds rate. The choice is arbitrary but does have the virtue of matching the preferences that have often been used in policy optimization experiments carried out for the FOMC; see Svensson and Tetlow (2005).

The results of this exercise can be summarized graphically. In Figure 10, the green solid line is the optimized coefficient on inflation, α_{Π}, while the blue dashed line is feedback coefficient on the output gap, α_Y. The response to inflation is universally low, never reaching the 0.5 of the traditional Taylor (1993) rule.[27] By and large, there is relatively little time variation in the inflation response coefficient. The output gap coefficient is another story. It too starts out low with the first vintage in July 1996 at about 0.2, but then rises almost steadily thereafter, reaching a peak of nearly 1 with the last vintage in November 2003. There is also a sharp jump in the gap coefficient over the first two quarters of 2001. One might be tempted to think that this is related to the jump in the sacrifice ratio, shown in Figure 6. In fact, the increase in the optimized gap coefficient precedes the jump in the sacrifice ratio, although the two might have common antecedents. The increase in the gap coefficient coincided with the inclusion of a new investment block in the model, which in conjunction with the supply block of the model, tightened the relationship between supply-side disturbances and subsequent effects on aggregate demand.[28] The new investment block, in turn, was driven by two factors: the addition by the Bureau of Economic Analysis a year earlier of software in the definition of equipment spending and the capital stock, and associated new appreciation, on the part of the staff, of the importance of the ongoing productivity and investment boom. In any case, while the upward jump in the gap coefficient stands out, it bears recognizing that the rise in the gap coefficient was a continual process.[29]

[27] That said, the measure of inflation differs here. In keeping with the tradition of inflation targeting countries, we use the rate of the change in the PCE price index as the inflation rate of interest. Taylor (1993) used the GDP price deflator.

[28] In essence, the linkage between a disturbance to total factor productivity and the desired capital stock in the future was clarified and strengthened so that an increase in TFP that may produce excess supply in the very short run can be expected to produce an investment-led period of excess demand later on.

[29] We conducted a similar exercise for so-called extended Taylor rules that include a term in the lagged federal funds rate. The results were essentially the same. In particular, the coefficient on the lagged fed funds rate was about 0.2 regardless of the vintage, and the coefficients on inflation and the output gap were slightly lower than in Figure 9, about enough to result in the same long-run elasticity. This result is consistent with the finding of Rudebusch (2001) for the Rudebusch-Svensson model, but differs from that of Williams (2003) for a linearized rational expectations version of the FRB/US model. The reason is that without rational expectations, the efficacy of "promising" future settings of the funds rate through instrument smoothing is impaired.

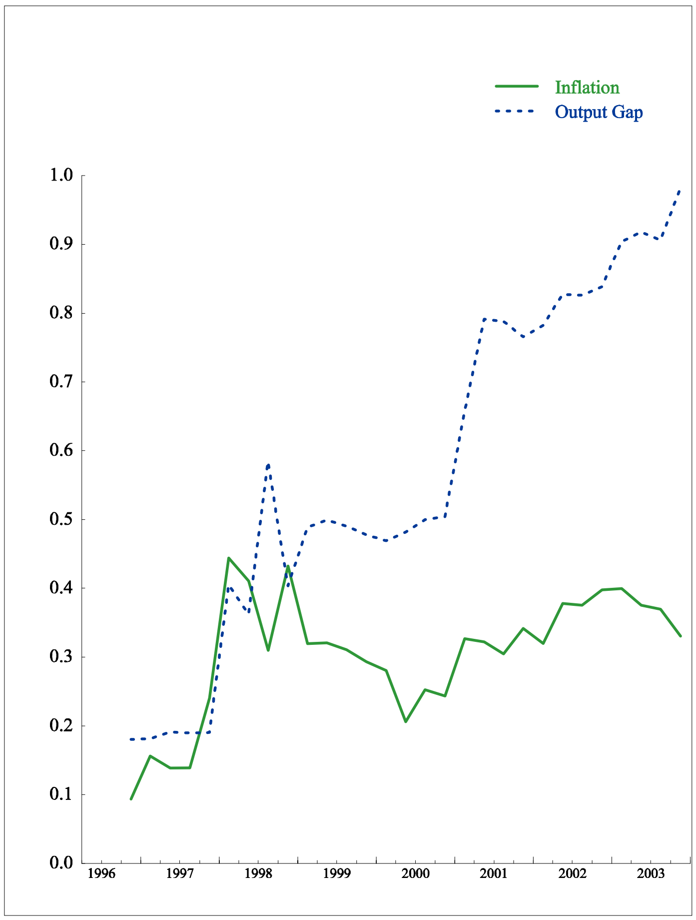


Figure 10: Optimized Taylor rule coefficients by vintage

4.2 *Ex post* optimal policies

Figure 10 showed the optimized Taylor rules. These policies were optimal given the parameterization of the rule and the average incidence of shocks as measured by the bootstrapped residuals. What is optimal, however, is always conditional on the maintained assumptions of the exercise. Most of this paper has been about model uncertainty. However uncertainty about the shocks one might be expected to face is also an issue. In an effort to examine this question, this subsection considers the Taylor-rule coefficients that would be optimal *had the Fed known the precise sequence of shocks* they were to experience over the 30 years prior to the vintage date. Whereas the coefficients in Figure 10 were chosen to minimize the loss function, (2), over bootstrapped draws of the residuals, here we pick the optimal coefficients to minimize the same loss function, over the same periods, but for the one sequence of draws that was actually experienced. In this way, Figure 10 can be thought of as the *ex ante* optimal coefficients, so called because those coefficients are optimal given that the Fed does not know the precise sequence of shocks, and here we will look at *ex post* optimal coefficients.

Obviously, the idea of an *ex post* optimal rule is an artificial concept. It assumes information that no one could have. Moreover, if one did have such information (and knew the model with certainty as well), it would not be reasonable to restrict oneself to a simple rule like the Taylor rule. Instead, one would choose precise values of the funds rate, period by period, to minimize the loss function. Our purpose here is to demonstrate the benefits of better information. Later on, we shall look at the other side of the coin by examining the costs of the hubris of believing too much. It is also worth noting that since the *ex post* optimal rules are conditional on just a single "draw" of shocks, they will tend to be sensitive to relatively small changes in specification or shocks and will vary a great deal from vintage to vintage.

The results are shown in Figure 11 and can be compared with those in Figure 10. It is worthwhile to divide the results into two parts, demarcated by vintage: the 1990s and the new century. Volatility aside, the *ex post* optimal output gap coefficients are mostly lower than the *ex ante* optimal ones, and the inflation coefficients are mostly higher. In the new century, the inflation and output gap coefficients rise more-or-less continuously in Figure 11. Thus the *ex ante*, *ex post* and estimated coefficients have a broadly similar pattern in the new century, albeit with differences in magnitude.

Of particular interest given the recent literature on the subject is the response to the output gap. Figure 12 shows the real-time output gap coefficients for the *ex ante* optimal coefficients (the green solid line) and the *ex post* optimal (the blue dashed line). In broad terms, the two lines share some features. Both are low (on average) in the early period; both climb steeply at the turn of the century, and both continue to climb thereafter, albeit more slowly. But there are interesting differences as well, with 1999 being an interesting period. This was a period where critics of the Fed argued that policy was too easy. The context was the three 25-basis-point cuts in the funds rate undertaken in 1998 in response to the Asia crisis and the Russian debt default. At the time, problems in the secondary market for sovereign debt had developed, resulting in thin trading. Worries about what this might signify for the global economy preoccupied Fed decision making. By 1999, however, these concerns were seen to have been overstated and so the FOMC starting "taking back" the previous decreases. To some, including Cecchetti *et al.* (2000) and Mussa (2003) the easier stance undertaken in late 1998 and into 1999 exacerbated the speculative stock market boom of that time and may have amplified the ensuing recession.[30] . The *ex post* optimal feedback on the output gap, shown by the blue dashed line, was volatile. For the 1999 models, and given the particular shocks over the period shown in the picture, the optimal response to the gap was zero; but within months, it rose to about 0.4. In contrast, the *ex ante* optimal coefficients were essentially unchanged over the same period, as were the more important multipliers, which indicates that changes in the shocks were critical. Given that the shock sets in 1999 and 2000 overlap, this is a noteworthy change. To us, the important point to take from this is not the proper stance of policy at that point in history, but rather that it is so dependent on seemingly small changes. Our analysis also hints at some advantages of discretion: the willingness to respond to the specific shocks of the day–if one is able to discern them. We shall have more to say about this a bit later.

5 Performance

To this point, we have compared model properties and the policies that those properties prescribe. Notwithstanding having specified a loss function, we have had nothing directly to say about performance. This section fills this void.

In the first subsection, we investigate how useful prior information about the sequence of shocks might be for policy and hence welfare. Specifically, we conduct counterfactual experiments on the single sequence of shocks immediately preceding each model vintage. Thus, this subsection is the performance counterpart to the design subsection of optimal *ex post* policies. It tells us the benefit of being right about the shock sequence underlying the *ex post* optimal policy. Then in subsection 5.2 , we consider the performance, on average of the model economies under stochastic simulation. The exercise in subsection 5.2 is a counterpart to the *ex ante* optimized policy rules in Figure 10. Among other things, it will tell us about the cost of being wrong in our beliefs about knowledge of the shocks.

5.1 Performance in retrospect: counterfactual experiments

If the *ex post* optimal rule really would have been optimal for each vintage of the model–conditional, of course, on that model–how much better would it have been

[30]In his memoir of his time on the Board of Governors, Lawrence Meyer considers this argument and mostly rejects it. See Meyer (2004), especially chapter 7.

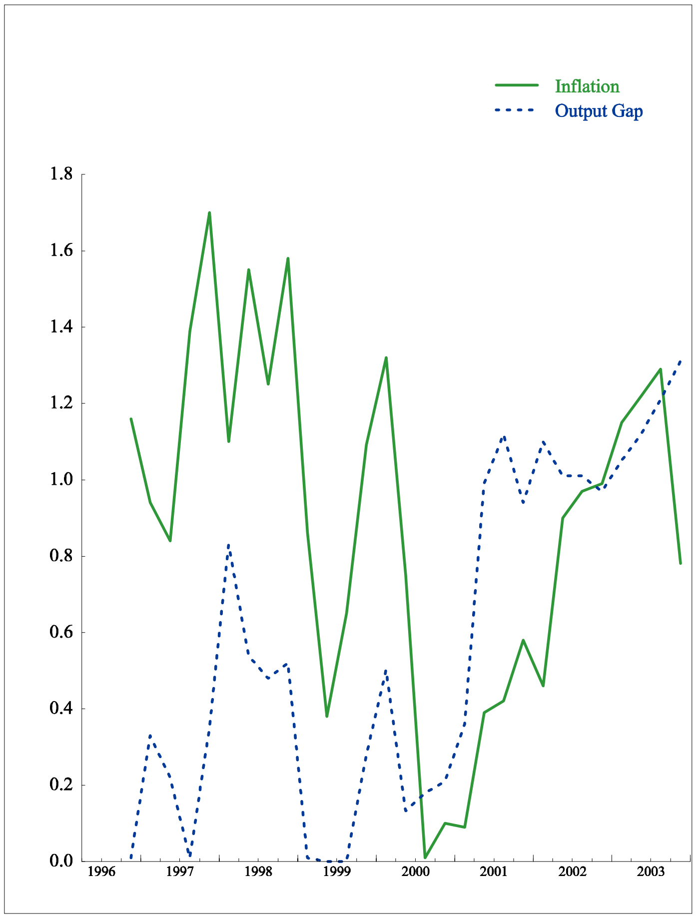


Figure 11: *Ex post* optimal Taylor rule coefficients by vintage

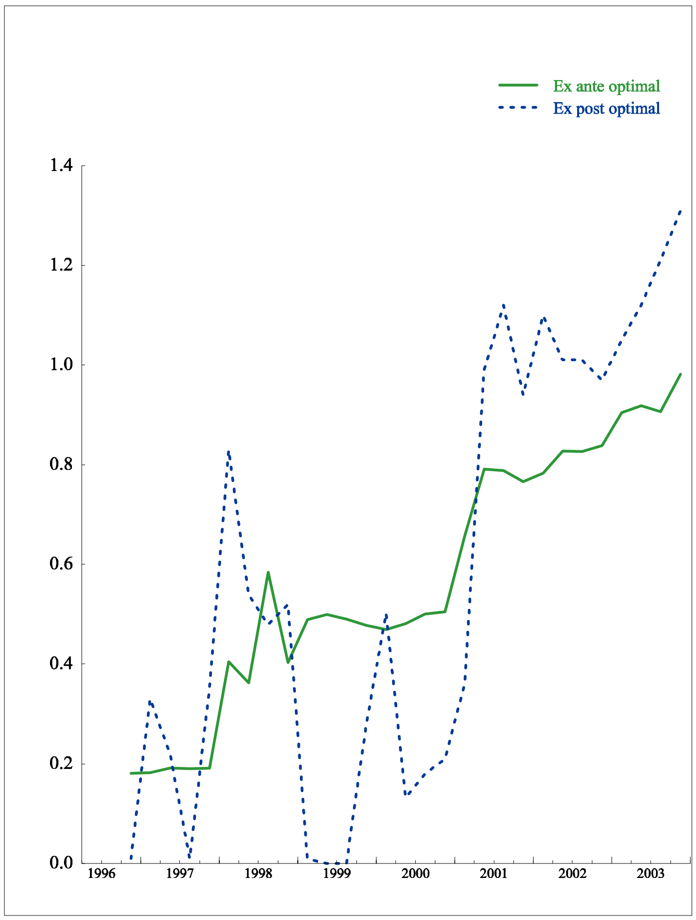


Figure 12: Comparison of output gap coefficients

than, say, the *ex ante* optimal rule? In other words, how valuable is that kind of information for the design of policy? We answer this question with a counterfactual simulation on selected model vintages. To facilitate comparison with the next subsection and still keep the size of the problem manageable, we restrict our attention to just two of our 30 model vintages. For this purpose, we elect to use the February 1997 and the November 2003 models. These were chosen because they were far apart in time, thereby reflecting as different views of the world as this environment allows, and because their properties are the most different of any in the set. In particular, the February 1997 model has the lowest sacrifice ratio of all vintages considered, and the November 2003 model has the highest. It follows that these two models should more-or-less encompass the results of other vintages.

The details of our simulation are straightforward: each simulation is initialized with the conditions as of 20 years and two quarters before the vintage itself, as measured by that model vintage. The simulation ends two quarters before the vintage date. Then the policy rule in question is allowed to control the economy without error subjecting the model to those shocks that the economy bore over the period. The loss in each instance is measured using the same loss function as in the optimization exercises, (2). The losses are then normalized such that the historical path represents a loss of unity. All other losses can be interpreted in terms of percentage deviations from the baseline loss.

The results are shown in Table 2 below. Let us focus for the time being on the left-hand panel with the results for the February 1997 model. To aid in the interpretation of the results, the policy rule's coefficients are shown, where applicable. According to the model, the *ex post* optimal policy would have been superior to the historical policy. This might seem obvious, since the *ex post* optimal policy has the benefit of "seeing" the shocks before they occur, but recall that the *ex post* optimal policy is constrained to respond to just the output gap and the inflation rate, whereas the FOMC in history faced no such constraint. Thus it is a noteworthy point that the *ex post* policy does better–almost twice as well–as the historical policy.[31] Interestingly, the traditional Taylor rule also does better than the historical policy. By contrast, the *ex ante* optimal policy does a fair amount worse. What both the Taylor rule and the *ex post* optimal policy share is stronger responses in general, and to inflation in particular, than the *ex ante* optimal policy. Evidently, the average sequence of shocks that conditions the *ex ante* optimal policy was less inflationary than the actual sequence.

[31] That said, as we noted before, the performance comparison assumes preferences that may not match the FOMC's preferences, although they are arguably very reasonable preferences.

Table 2

Normalized model performance in counterfactual simulation*

	February 1997 vintage			November 2003 vintage		
	α_π	α_y	$L\alpha$	$_\pi$	α_y	L
Historical policy	-	-	1	-	-	1
Ex post optimal	0.94	0.33	0.56	0.78	1.31	2.25
Ex ante optimal	0.18	0.25	1.80	0.30	1.07	4.17
Taylor rule	0.50	0.50	0.74	0.50	0.50	10.79

* Selected rules and model vintages. Using the estimated shocks over 20 years.

The right-hand panel shows the results for the November 2003 vintage of the model. Here the results are much different. The historical policy is substantially better than any of the alternative candidates. This suggests that responding to just two variables is insufficient for the shocks borne during this period. If the best two coefficients of the *ex post* optimal policy were less than ideal, the basic Taylor rule and the *ex ante* policy should do worse, and indeed they do: much worse. The lower the feedback on the output gap in these scenarios, the poorer the performance. The reasons for this should not be surprising: the shocks during this period included shocks to the growth rate of potential output, as outlined in Figure 3 above. Such shocks manifest themselves in more variables than just the output gap and inflation. Indeed, the short-run impact of an increase in productivity is to reduce inflation and raise output, leading to offsetting effects on policy. However as time goes by, the higher growth rate of productivity raises the desired capital stock thereby increasing the equilibrium real interest rate. The Taylor rule and its cousins are ill designed to handle such phenomena.

5.2 Performance on average: stochastic simulations

Another way that we can assess candidate policies is by conducting stochastic simulations of the various model vintages under the control of the candidate rules and evaluating the loss function. We do this here. We subject both of these models to same set of stochastic shocks as in the *ex ante* optimization exercise. Under these circumstances, the *ex ante* optimal rule must perform the best. Accordingly, in this case, we normalize the loss under the *ex ante* optimal policy to unity. The results are shown in Table 3.

Table 3

Normalized model performance under stochastic simulation*

	February 1997 vintage			November 2003 vintage		
	α_π	α_y	$L\alpha$	π	α_y	L
Ex ante optimal	0.18	0.25	1	0.30	1.07	1
Ex post optimal	0.94	0.33	1.76	0.78	1.31	4.19
Taylor rule	0.50	0.50	1.33	0.50	0.50	1.49

* Selected rules and model vintages. 400 draws of 80 periods each.

For the moment, let us focus on the left-hand panel, with the results for the February 1997 model; once again, we show the coefficients of the candidate rules for easy reference. The *ex ante* optimal coefficients are both low, at about 0.2. The *ex post* optimal coefficients are higher, particularly for inflation. However, the table shows that applying the policy that was optimal for the particular sequence of shocks to the average sequence, selected from the same set of shocks, would have been somewhat injurious to policy performance, with a loss that is 76 percent higher. The Taylor rule prescribes stronger feedback on output but weaker feedback on inflation, than the *ex ante* optimal policy. The fact that the loss under the Taylor rule is approximately midway between that of the *ex ante* and *ex post* rules suggests that it is the response to inflation that is the key to performance for this model vintage and the corresponding shock set. Still, in broad terms, none of the rules considered here performs too badly for this vintage.

The results for the November 2003 vintage, shown in the right-hand panel, are in some ways more interesting. Recall that in Table 2 we showed that the *ex post* optimal rule performed approximately twice as well as the *ex ante* optimal rule for the particular sequence of shocks studied. Here it is shown that this same *ex post* optimal rule–that is optimal for the specific shocks *in the particular order* of the period immediately before the vintage–performs very poorly for the same shocks *on average*. The reasons are clear from our prior examinations. The period ending in mid-2003 contained a number of important, correlated shocks; namely, the productivity boom and the stock market boom. The episodic nature of these disturbances makes them special. With knowledge of these shocks including the order of their arrival, a policy–even a policy constrained to respond to just two objects, inflation and the output gap–can be devised to do a reasonable job. But with randomization over these shocks, so that one knows their nature but not the specific order, the best policy is very different. This tells us is about the cost of hubris: a policy maker that thinks he knows a lot about the economy and acts on that belief, may pay a substantial price if the world turns out to be different than he expected. This impression is amplified by the Taylor rule which show performances that, while inferior to the *ex ante* optimal rule–as they must be–are not too bad.

One might wonder why the November 2003 model is so much more sensitive to policy settings than the February 1997 model. Earlier, we noted that performance in general is jointly determined by initial conditions (that is, the baseline), the stochastic shocks, the model and the policy rule. All of these factors are in play in these results. However, as we indicated in the previous subsection, the nature of the shocks is an important factor. The shocks for the February 1997 model come from the relatively placid period of the late 1960s to the mid-1990s, whereas the shocks to the November 2003 model contain the disturbances from the mid-1990s. We tested the importance of these shocks by repeating the experiment in this subsection using the November 2003 but restricting the shocks to the same range used for the February 1997 vintage. Performance was markedly better regardless of the policy rule. Moreover, there was less variation in performance across policy rule specifications. Since, however, the stochastic shocks come from the same data that render the model respecifications, this just emphasizes the importance of model uncertainty in general, and designing monetary policy to respond to seemingly unusual events in particular.

6 Concluding remarks

This paper has examined real-time model uncertainty in the United States. To do this, we have exploited an archive of every vintage of the FRB/US model of the macro economy since the model's inception as the Board of Governor's macroeconometric model in 1996. We examined how the model properties have changed over time and how the optimal policies for those vintages have changed alongside.

We found that the time variation in model properties is surprisingly substantial. Surprising because the period under study, at eight years, is short; substantial because the differences in model properties over time imply large differences in optimized policy coefficients.

We also compared different policies by model vintage, doing so in two different ways. In one rendition, we compared policies conditional on bootstrapped model residuals; in the other, we conducted counterfactual simulations examining performance over approximately the same period where the model vintage was estimated. Besides finding that our optimized rules differ by vintage, we also found that plausible alternatives to the optimized policy result in significant incremental losses. This puts policy makers in the horns of a dilemma. On the one hand, time variation in the FRB/US model suggests that model uncertainty is substantial and thus policies should be designed with an eye toward minimizing the implications of misspecification of the model. On the other hand, the results also suggest that performance depends in an important way on the particular characteristics of the rule.

Finally, we found conflicting evidence on the efficacy of discretion in monetary policy, defined here as the FOMC picking the federal funds rate period by period, depending on the conditions of the day. In one of the two models we studied, the historical policy was better than even the best policy rule. In another case, however, the best two-parameter rule performs better than the historical path in spite of the fact that the historical policy can react to a much wider range of variables.

We have shown that model uncertainty matters for policy. But none of the policies considered here explicitly considered model uncertainty in their design. A useful extension would be to examine the extent to which common methods for dealing with uncertainties would ameliorate the problems identified here. This subject is taken up in Tetlow (2004b).

References

[1] Benigno, P. and Woodford M. (2004) "Linear quadratic approximation to optimal policy problems" unpublished manuscript delivered to the 10th International Conference on Computing in Economics and Finance, Society for Computational Economics, University of Amsterdam, July 8-10,2004.

[2] Bernanke, B., Gertler, M. and Gilchrist, S. (1999) "Monetary policy and asset price volatility" in *New Challenges for Monetary Policy* (Kansas City: Federal Reserve Bank of Kansas City): 77-1229

[3] Bernanke, B. and Gertler, M. (2001) "Should central banks respond to movements in asset prices?" *American Economic Review*,91(2): 253-257.

[4] Blanchard, O. and Khan, C. (1980) "The solution of linear difference equations under rational expectations" *Econometrica*,48: 1305-1311.

[5] Brainard, W. (1967) "Uncertainty and the effectiveness of monetary policy" *American Economic Review* 57(2): 411-425.

[6] Brayton, F. and Tinsley, P (eds.) "A Guide to FRB/US – a macroeconomic model of the United States" Finance and Economics Discussion Series paper no. 1996-42, Board of Governors of the Federal Reserve System.

[7] Brayton, F., Mauskopf, E., Reifschneider, D., Tinsley, P., and Williams, J.C. (1997) "The role of expectations in the FRB/US macroeconomic model" *Federal Reserve Bulletin* (April), pp. 227-245.

[8] Brayton, F., Levin, A.T., Tryon, R., and Williams, J.C. (1997) "The evolution of macro models at the Federal Reserve Board" *Carnegie-Rochester Conference Series on Public Policy*,47: 227-245.

[9] Cecchetti, S., Genberg, H., Lipsky, J., and Wadhwani, S. (2000) *Asset Prices and Central Bank Policy* The Geneva Report on the World Economy, vol. 2 (London: Center for Economic Policy Research).

[10] Christiano, L.J., Eichenbaum, M., and Evans, C. (2004) "Nominal rigidities and the dynamic effects of a shock to monetary policy" *Journal of Political Economy.*(in press).

[11] Cogley, T. and Sargent, T. (2004) "The conquest of U.S. inflation: learning and robustness to model uncertainty" unpublished manuscript, University of California at Davis and New York University (October).

[12] Croushore, D and Stark, T. (2001) "A real-time data set for macroeconomists" *Journal of Econometrics*,105(1): 111-130.

[13] De Long, B. (1997) "America's peacetime inflation: the 1970s" in C. Romer and D. Romer (eds.) *Reducing Inflation: motivation and strategies* (Chicago: University of Chicago Press): 247-276.

[14] Giannoni, M.P. (2002) "Does model uncertainty justify caution?: robust optimal monetary policy in a forward-looking model" *Macroeconomic Dynamics*,6(1): 111-144.

[15] Hansen, L.P. and Sargent, T.J. (1995) "Discounted linear exponential quadratic Gaussian control" *IEEE Transactions on Automatic Control,*40: 968-971.

[16] Hansen, L.P. and Sargent, T.J. (2004) *Misspecification in Recursive Economic Theory* (unpublished monograph, November 2003)

[17] Hansen, L. and Sargent, T. (2003) "Robust control of forward-looking models" *Journal of Monetary Economics* 50(3): 581-604.

[18] Henderson, D.W. and McKibbin, W.J. (1993) "A comparison of some basic monetary policy regimes for open economies: implications of different degrees of instrument adjustment and wage persistence" *Carnegie-Rochester Conference Series on Public Policy,*39: 221-317.

[19] Kim, J. and Kim, S.H. (2003) "Spurious welfare reversals in international business cycle models" *Journal of International Economics*,60(2): 471-500.

[20] Levin, A., Wieland, V. and Williams, J.C. (1999) "Monetary policy rules under model uncertainty" in J.B. Taylor (ed.) *Monetary Policy Rules* (Chicago: University of Chicago Press).

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

[22] McCallum, B. (1988) "Robustness properties of a rule for monetary policy" *Carnegie-Rochester Conference Series on Public Policy*,29: 173-204.

[23] Meyer, L. (2004) "A term on the Fed: an insider's view" (New York: HarperCollins).

[24] Mussa, M. (2003) "Asset prices and monetary policy" in: W.C. Hunter, G.G. Kaufman, and M. Pomerleano (eds.) *Asset Price Bubbles: the implications for monetary, regulatory, and international policies* (Cambridge: MIT Press): 40-50.

[25] Onatski, A (2003) "Robust monetary policy under model uncertainty: incorporating rational expectations" unpublished manuscript, Columbia University.

[26] Orphanides, A. (2001) " Monetary policy based on real-time data" *American Economic Review*,91(4): 964-985.

[27] Orphanides, A, Porter, R, Reifschneider, D, Tetlow, R, and Finan, F. (2000) "Errors in the Measurement of the Output Gap and the Design of Monetary Policy" *Journal of Economics and Business*,52: 117-141.

[28] Roberts, J. M. (2004) "Monetary policy and inflation dynamics" Finance and Economics Discussion Series paper no. 2004-62, Board of Governors of the Federal Reserve System.

[29] Romer, C. and Romer, D. (2000) "Federal Reserve information and the behavior of interest rates " *American Economic Review*,93(3): 429-457.

[30] Romer, C. and Romer, D. (2002) "The evolution of economic understanding and post-war stabilization policy" National Bureau of Economic Research working paper no. 9274.

[31] Rotemberg, J and Woodford, M. (1997) "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy" *National Bureau of Economic Research Macroeconomics Annual*, 1997: 297-346.

[32] Rudebusch, G. (2001) "Is the Fed too timid?: monetary policy in an uncertain world" *Review of Economics and Statistics*,83(3): 203-217.

[33] Sack, B. (1999) "Does the fed act gradually: a VAR analysis" *Journal of Monetary Economics*,46(1): 229-256.

[34] Sack, B. and Wieland, V. (2000) "Interest-rate smoothing and optimal monetary policy: a review of recent empirical evidence" *Journal of Economics and Business*,52(1/2): 205-228.

[35] Soderstrom, U. (2002) "Monetary policy with uncertain parameters" *Scandinavian Journal of Economics*,104(1): 125-145.

[36] Svensson, L.E.O. (1999) "Inflation targeting as a monetary policy rule" *Journal of Monetary Economics*,43(3): 607-654.

[37] Svensson, L.E.O. (2002) "Inflation targeting: should it be modeled as an instrument rule or a targeting rule?" *European Economic Review*,46(4/5): 771-180.

[38] Svensson, L.E.O. and Tetlow, R. (2005) "Optimum policy projections" unpublished manuscript (February) http://www.cox.members.net/btetlow/default.htm.

[39] Svensson, L.E.O. and Woodford, M. (2003) "Optimal indicators for monetary policy" *Monetary Economics,*46(1): 229-256.

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

[41] Tetlow, R. (2004a) "Monetary policy, asset prices and misspecification: the robust approach to model uncertainty and bubbles" unpublished manuscript, Division of Research and Statistics, Federal Reserve Board.

[42] Tetlow, R. (2004b) "Real-time model uncertainty in the United States: robust policies" unpublished manuscript in progress, Division of Research and Statistics, Board of Governors of the Federal Reserve System.

[43] Tetlow. R. and von zur Muehlen, P. (2001) "Robust monetary policy with misspecified models: does model uncertainty always call for attenuated policy?" *Journal of Economic Dynamics and Control* 25(6/7): 911-949.

[44] Tetlow, R. and von zur Muehlen, P. (2004) "Avoiding Nash Inflation: Bayesian and robust responses to model uncertainty" *Review of Economic Dynamics*,7(4): 869-899.

[45] Williams, J. C. (2003) "Simple rules for monetary policy" *Federal Reserve Bank of San Francisco Economic Review:* 1-13.