

Dynamic Limited Dependent Variable Modeling and U.S.

Monetary Policy

George Monokroussos*

Department of Economics

University of California, San Diego

November 2004

JOB MARKET PAPER

Abstract

I estimate, using real-time data, a forward-looking monetary policy reaction function that is dynamic and that also accounts for the fact that there are substantial restrictions in the period-to-period changes of the Fed's policy instrument. I find a substantial contrast between the periods before and after Paul Volcker's appointment as Fed Chairman in 1979, both in terms of the Fed's response to expected inflation and in terms of its response to the (perceived) output gap: In the pre-Volcker era the Fed's response to inflation was substantially weaker than in the Volcker-Greenspan era; conversely, the Fed seems to have been more responsive to real activity in the pre-Volcker era than later.

*I am indebted to James Hamilton for his advice and encouragement, and for many useful discussions. I am also grateful to Graham Elliott and to Allan Timmermann for their feedback and for useful discussions. I thank Fabio Buseti, Carlos Capistrán, Masako Miyanishi, Davide Pettenuzzo, Valerie Ramey, Yixiao Sun, and participants in the UCSD Econometrics seminar for useful comments. Finally, I am grateful to Jean Boivin and to Athanasios Orphanides for generously providing me with their real-time data - the unemployment and inflation data for the early years used in this study come from Boivin's data set. Any errors are mine. Please send comments to gmonokro@ucsd.edu. Updated versions of the paper will be made available at: <http://www.econ.ucsd.edu/~gmonokro/>

1 Introduction

One question that has received much attention in recent work on US monetary policy is that of the contrast in US macroeconomic performance between the periods before and after the appointment of Paul Volcker as chairman of the Federal Reserve in August of 1979, and of the causes of that contrast. As Figure 1 demonstrates, the late 1960s and the 1970s were a period of relative macroeconomic instability, as they were characterized by high and volatile rates of inflation. This era of the *Great Inflation* contrasts starkly however with what followed after Paul Volcker's appointment, namely a period of low inflation and stable output growth.

Understanding the causes of this dramatic change is essential both for a fair historical assessment of past policies and for the design of a better monetary policy for the future. This issue has thus justifiably been scrutinized in numerous studies, and various explanations have been proposed. One influential line of research focuses mainly on intertemporal differences in monetary policy and emphasizes the role that improved policies in the Volcker-Greenspan eras played in achieving better macroeconomic outcomes during the 1980s and the 1990s.

This literature that focuses mostly on the role of monetary policy makes extensive use of single equation reaction functions for the Federal Reserve, also known as Taylor rules, that link the Fed's policy instrument (typically taken to be monthly or quarterly averages of the federal funds rate - the interest rate that banks charge each other for overnight loans of Federal Reserve deposits) to measures of the inflation gap and of the real output gap¹ (or of the unemployment gap). Such reaction functions are compatible with broadly held views about what the central goals of US monetary policy are, and appropriately chosen rules from that family of reaction functions can have both a normative and a positive justification (see, inter alia, Taylor (1993), Judd and Rudebusch (1998), and Clarida Galí and Gertler (1999)).

One of the better known papers from that literature that uses reaction functions is Clarida, Galí, and Gertler (2000) (henceforth CGG). CGG take a standard Taylor rule and augment it by introducing dynamics, in the form of lags of the dependent variable as additional explanatory

¹The inflation and output gaps are gaps of these variables from their desired target levels.

variables, and they also employ a forward-looking framework. Both of these arguably capture salient features of the Federal Reserve's behavior:

The lags of the policy instruments allow for a certain degree of inertia and for possible partial adjustment, which is compatible with the widely held view that the Fed tends to smooth changes in interest rates - something also evident in the data. In particular, the Fed tends to adjust its target for the federal funds rate in a cautious manner, and only gradually, with slow, stodgy steps, at discrete points in time. Reasons for this inertia in monetary policy include data uncertainty, and more generally uncertainty about the structure of the economy. Indeed, Fed actions, taken on the basis of possibly inaccurate real-time information, that prove ex-post to have been too aggressive and to have been taken too hastily would compromise the monetary authority's credibility.

The forward looking framework is one where *forecasts* of inflation and the output gap for one or more periods ahead are employed in the reaction function, rather than contemporaneous or lagged values for these variables (of a backward looking specification). Thus, the forward looking specification essentially implies that the Fed considers a broader information set in determining the target level for its policy instrument, rather than being restricted to considering only contemporaneous or past values of the output and inflation gaps, and it also gives a more realistic description of how sophisticated monetary policy ought to look like in the face of likely lags in the monetary transmission mechanism.

CGG estimate such a dynamic and forward-looking reaction function for the postwar U.S. economy. Their central finding is that monetary policy prior to Volcker's appointment *accommodated inflation* - the Fed typically raised the nominal interest rates by less than the increase in expected inflation, which would thus result in a lower real interest rate², while in the Volcker-Greenspan era the Fed drastically changed its approach and adopted a much more anti-inflationary stance, raising not only nominal, but also real interest rates in response to increases in expected inflation; it thus contributed, at a very minimum, to the transition from the volatile 1970s to the rosy 1980s and 1990s - a period of stability and low inflation.

²Thus, the Pre-Volcker Fed did not satisfy *Taylor's principle*.

This is perhaps not a surprising finding to many - indeed the notion that Paul Volcker and Alan Greenspan did a superior job in conducting monetary policy than their predecessors may sound by now more like conventional wisdom than like a controversial proposition. Nevertheless this central empirical finding of CGG of a stark, dramatic contrast between the pre- and post- Volcker conduct of monetary policy has been at the center of much recent attention (probably justifiably given the importance of the question) and has been complemented, refined, or even challenged on several grounds in subsequent research.

Related Past Literature

First, there is another influential subset of the recent literature on U.S. monetary policy that focuses more on shocks and on changes in their volatilities, rather than on changes in monetary policy. This literature (see, inter alia, Kim and Nelson (1999b), Sims (1999), Blanchard and Simon (2001), Stock and Watson (2003), and Sims and Zha (2004)) emphasizes the role of non-policy shocks, whose volatility was higher in the pre-Volcker era than in the Volcker-Greenspan era, and it argues that it was this, rather than any policy changes, that was the central factor behind the observed dramatic change in macroeconomic outcomes of the two periods.

Yet another interesting line of recent work argues that changes in U.S. monetary policy were more gradual than suggested by CGG or that there were richer dynamics than suggested by the simple split-sample approach and concentrates on estimating reaction functions using Time Varying Parameter approaches. Papers along these lines include Cogley and Sargent (2001, 2002), Boivin (2004), and Jalil (2004).

Furthermore, and as Orphanides demonstrates in his influential 2001 paper, it is essential, when trying to reach policy conclusions on the basis of estimated reaction functions, to use real-time data, that is data that were actually available to the Federal Open Markets Committee at the time their decisions were made, rather than revised data series (typically used in applied work prior to Orphanides' contribution) that became available only ex post.

Orphanides extends this approach further in subsequent papers (such as Orphanides (2002,

2004)), and finds, using such real-time data and forward-looking reaction functions that, in contrast to the CGG conclusions, monetary policy in the pre-Volcker era, far from being accommodative to inflation, was in a similar manner to the Volcker-Greenspan era, activist, forward-looking, and strong and decisive in its reaction to inflationary surges. While this contradicts the CGG conclusion that the instability of the 1970s was at least in part due to weak monetary policy, Orphanides also finds that monetary policy prior to Volcker's appointment was "too activist in reacting to perceived output gaps that retrospectively proved overambitious"³. That is, the Fed's real-time estimates of potential output were upwardly biased, and this resulted in the Fed responding to these biased estimates by following policies that in retrospect, and using the knowledge of the revised and corrected estimates of the output gap, would be judged to have been too expansionary, which thus probably contributed to the instability and inflationary pressures of that period.

All the contributions outlined above are not mutually exclusive and should be taken into account in careful attempts to assess the historical evolution of U.S. monetary policy and of its impact on the economy. However, one issue which, while equally important, has not been given nearly as much attention in the literature is that of characterizing the Fed's policy instrument, and specifically, of determining what exactly the nature of the policy instrument is, and of taking into account the specifics of its time series behavior in estimation exercises of reaction functions.

Characterizing the Fed's Policy Instrument

As suggested above, the obvious candidate for the Fed's policy instrument is the federal funds rate - indeed that has been the presumed policy instrument in most of the studies that investigate U.S. monetary policy with Taylor rules, including the ones outlined above. While there is no explicit institutional directive that defines the federal funds rate (or any other variable) to be the Fed's policy instrument, related empirical work, including Bernanke and Blinder (1992) and Bernanke and Mihov (1998) establishes that, with a possible exception of non-borrowed reserve targeting for a brief period during the first three years of Paul Volcker's tenure, the Fed has indeed treated the

³Orphanides (2004).

fed funds rate as its policy instrument⁴.

Further evidence to this is provided by the fact that the Fed has been explicitly announcing a *target* for the fed funds rate (henceforth the target) since 1994, usually during Federal Open Market Committee meetings, but sometimes in between such meetings too. Even before 1994 changes in the target were quite accurately inferred by the market. Indeed, as Hamilton and Jordà (2002) note, the Wall Street Journal would report any movements of the target on a daily basis, and related academic literature has compiled time series of target changes for periods prior to 1994 (such as Cook and Hahn (1989) for the mid- and late 1970's and Rudebusch (1995) for the mid-1980's and later)⁵. Announcements of changes or no changes to the target, and in general news related to the target and its movements has always been the subject of intense interest by the markets, *precisely because the fed funds rate and its target are understood to be the policy instrument of the Federal Reserve*.

However, there are quite severe restrictions in the way the target changes from period to period: At any given period the target either will not change at all or it will change, by multiples of 25 basis points since November of 1989 and by multiples of 6.25 basis points earlier. Thus, the changes in the target fall into a small number of discrete categories. This discreteness is preserved when monthly averages of the target⁶ are considered too: For example, for the Greenspan period⁷, the changes in monthly averages of the target are equal to 0 for 42.4% of the time.

Essentially the same holds true for periodic averages of the federal funds rate: Once the target has been determined, it is then the task of the Federal Reserve Bank of New York, and of its trading desk in particular, to conduct open market operations (purchases or sales of Treasury securities), based on careful monitoring of commercial banks' reserve requirements and available Fed deposits,

⁴Even for these three years at the start of Volcker's term, it is being argued (see, for example, Goodfriend (1991)) that the Fed had an implicit target for the federal funds rate.

⁵For further details on the mechanics, history, and stylized facts of setting the target one can read, inter alia, Meulendyke (1998), Hamilton and Jordà (2002), Piazzesi (2001), and Dueker (1999).

⁶The monthly averages of the target are taken at the "FOMC frequency", that is, over months during which there was an Federal Open Markets Committee meeting. The reason is that it is only during those months that Greenbook forecasts (that are used to construct the explanatory variables of the reaction functions estimated in this paper) were available. Further details on the data used are provided in Section 3.

⁷The sample used in this study ends in June of 1998.

so as to implement the set target. In other words, the New York Fed essentially acts as the Fed's broker, and ensures that the fed funds rate never fluctuates too far away from the set target. This fact is further illustrated in Figure 2, which plots the time series of monthly averages of the effective fed funds rate, of monthly averages of the target, and of end-of-month values of the target and shows how these three series are almost indistinguishable from each other.

Contributions of this Paper

It is thus evident that there are severe restrictions in the way the Fed's policy instrument changes from period to period, and so these restrictions ought to be taken into account in modeling and estimation exercises of Fed reaction functions. Estimating linear reaction functions with Gaussian error terms ignores these restrictions to the support space of the dependent variable and may thus lead to serious biases. However, little work has been done in that direction. All of the literature outlined above, and nearly all of the rest of the literature in the area employs linear specifications with Gaussian error terms. One exception to this paradigm is Hamilton and Jordà (2002), who propose the *Autoregressive Conditional Hazard* (ACH) approach which allows them to model the target as a discrete time series variable. However, this approach is more geared towards forecasting and falls outside the Taylor rule framework of all the studies that have been outlined above. Also, Dueker (1999a) takes these restrictions into account as he models the reaction function using a multinomial ordered probit for the changes of the target, (which is the approach adopted here too). However he does not account for possible serial correlation in the residuals, which is likely to be present.

In general, most existing applied macroeconomic work, including the work on monetary policy reaction functions, either ignores the restrictions to the support space of the dependent variable under consideration and focuses on the needed time series modeling requirements, or, conversely, employs Limited Dependent Variable estimation techniques at the expense of time series modeling. The main reason for that is that incorporating both of these at the same time results in an estimation task that presents the researcher with formidable computational challenges because of the need for

integration of multiple integrals with no closed form solution whose dimensionality can be the same as the time series dimension of the data, and/or because of the need for numerical optimization of difficult objective functions.

The *econometric contribution* of this paper is that I propose a flexible and practical algorithmic framework that overcomes the multiple integral problem by using *Markov Chain Monte Carlo* (henceforth MCMC) simulation techniques with *data augmentation*, whereby simulated samples of the latent variables typically used to model probits are generated from their model-implied conditional distributions. I propose a new and computationally attractive way of implementing the required smoothing algorithm for the generation of the latent variables that relies on the joint distributions of the latent residuals.

An added advantage of the MCMC approach and of the *Gibbs sampler* that is employed in this paper in particular is that additional time series features, such as Regime Switching, or Time Varying Parameters, that would substantially complicate the needed computational tasks in an extremum framework, can be easily introduced in the MCMC context as additional *blocks* in the Gibbs sampler, and this paper illustrates how this is done.

Finally, I overcome the additional computational challenges associated with difficult numerical optimizations by proposing a Bayes estimator that is a statistic of the simulated posterior distribution of interest, and which thus falls outside the extremum framework.

The *macroeconomic contribution* of the paper is that, in contrast to past literature, I model the Fed's reaction function in a way that does not ignore the discrete nature of the changes in the Fed's policy instrument, while also taking into account contributions of past literature, such as the ones outlined earlier. Specifically, I estimate, using real-time data compiled from Greenbook forecasts of the Federal Reserve, a forward-looking, dynamic ordered probit reaction function, and I also consider a series of extensions and robustness checks.

My central finding is that, while there is some evidence that the Fed was more activist in its response to real-time estimates of the output gap in the pre-Volcker era than later, there is also strong evidence that it was much less aggressive in its response to inflation in the pre-

Volcker era than later. Thus, while the Fed’s possibly excessive responsiveness to inaccurate real-time estimates of the output gap probably contributed to the instability of the 1970s, it is apparent that, as CGG claim, the Pre-Volcker Fed’s less than aggressive stance in fighting inflation, which changed dramatically with Volcker’s appointment, was a central factor behind the observed contrast in macroeconomic performance between the 1970s and the 1980s-1990s.

The plan for the rest of the paper is as follows: Section 2 develops the benchmark model and describes the associated estimation challenges, as well as the proposed estimation strategy that overcomes these challenges. Section 3 describes the data used, Section 4 provides a discussion of the estimation results of the benchmark model, while Section 5 considers a series of robustness checks and extensions, and Section 6 provides concluding remarks. The details of the algorithms employed in this paper are included in the Appendix.

2 A Forward Looking, Dynamic Ordered Probit Reaction Function for the Federal Reserve

A useful starting point is Taylor’s original specification for the reaction function of the Fed. Taylor (1993) suggested a very specific and simple linear rule:

$$ff_t = \pi_t + ff^N + 0.5(\pi_t - \pi^*) + 0.5y_t^{gap}, \quad (1)$$

where ff_t is the Fed’s policy instrument, the federal funds rate, π_t is the rate of inflation, ff^N is a natural real fed funds rate (that is, an equilibrium real fed funds rate that is consistent with full employment), $(\pi_t - \pi^*)$ is the inflation gap (that is, actual inflation minus a target inflation rate), and y_t^{gap} is the output gap (the difference between real output and potential output). As we can see from equation (1), Taylor did not estimate his reaction function. Rather, he suggested this linear rule, with the assumed coefficients of 0.5, and assumed rates of target inflation and natural real fed funds rate that are both 2%, as a useful and simple way to capture salient aspects of the

way the Fed decides what the appropriate level for its policy instrument is at each period. Indeed, rules such as this one can have a normative motivation, and are justified both in terms of what the legislated goals of US monetary policy are, and have also been performing relatively well when compared to other alternatives that the literature has proposed. For instance, and as Judd and Rudebusch (1998) calculate for the rule of equation (1), the R^2 for the Greenspan period is 87% for quarterly levels of the nominal funds rate, and 52% for quarterly changes⁸. Taylor rules have been used with similar success for other countries as well (such as, for example, Clarida, Galí and Gertler (1998) who demonstrate the empirical relevance of Taylor rules for a set of European countries).

The literature on reaction functions that followed Taylor's 1993 paper expanded the set of specifications considered and also generally took up the task of estimating such reaction functions. As discussed earlier, two useful features that much recent work has incorporated into reaction function specifications are a forward looking framework and dynamics.

A forward-looking specification means that the Fed sets its policy instrument in response to its expectations about future⁹ values of the inflation and output gap variables. This is arguably a more realistic depiction of the monetary authority's behavior in the face of lags in the monetary transmission mechanism that are generally present, and also it encompasses backward looking specifications, in the sense that a forward-looking Fed considers a broad information set in determining its forecasts for the inflation and output gaps, which includes past and present values of these variables.

In contrast to the specification of equation (1), which makes the restrictive assumption of an immediate adjustment of the fed funds rate to the level dictated by the inflation and output gap variables, a dynamic specification is again a more realistic depiction of observed Fed behavior as it captures the Fed's well documented tendency to smooth interest rates¹⁰, which can also be easily seen, for instance, in the actual time series for the fed funds rate, and can be obtained simply by introducing lags of the policy instrument¹¹ as additional explanatory variables¹¹.

⁸However, this simple specification isn't as successful in describing the Burns and the Volcker periods.

⁹Usually for a short forecast horizon.

¹⁰See, inter alia, Clarida, Galí and Gertler (1999), Taylor (1999), Sack and Wieland (2000), and Piazzesi (2001).

¹¹Rotemberg and Woodford ((1997), and (1999) - available in Taylor (1999)) offer an interesting motivation for the lagged dependence aspect of interest rate smoothing practiced by the Fed, arguing that a dynamic reaction function gives the Fed increased leverage over the long term interest rates , since the Fed can manipulate such long term rates

Such a dynamic, forward looking reaction function for the Fed's *desired level* for the federal funds rate at time t , namely ff_t^* is given by (2):

$$ff_t^* = \alpha + \rho(L)ff_{t-1}^* + \beta\pi_{t,h} + \gamma y_{t,h}^{gap} + \varepsilon_t^*, \quad (2)$$

where $\rho(L) = \rho_1 + \rho_2L + \dots + \rho_nL^{n-1}$, and where all of the roots of the associated polynomial $1 - \rho_1L - \rho_2L^2 - \dots - \rho_nL^n$ lie outside the unit circle¹². $\varepsilon_t^* \sim N(0, \sigma^2)$, $\pi_{t,h}$ is the h -period ahead forecast of inflation that the Fed made at time t , and $y_{t,h}^{gap}$ is the h -period ahead forecast of the output gap that the Fed made at time t . Thus, both of these are real-time forecasts, the definitions and details of which are provided in the data section that follows. Finally, and considering (2) in the context of Taylor rules, the intercept term α is modeled to capture both the natural real fed funds rate, and the target inflation rate, which are therefore not separately identifiable.

However, and as argued in the introduction, a drawback of the linear specification of equation (2) is that it neglects the discrete nature of the period-to-period movements of the Fed's policy instrument and thus that it may be confounding important aspects of the monetary policy making process: The target does not change at all for about half the time, and when it does change, it does so by multiples of 25 basis points (since November of 1989 and by multiples of 6.25 basis points earlier). Thus, the period-to-period changes of the target for the fed funds rate have historically fallen into a small number of categories¹³, and thus there is a clear distinction between this variable and the dependent variable modeled by equation (2): Equation (2) is a linear specification with a Gaussian error term, and it thus implies that there are no restrictions in the period-to-period movements of its dependent variable.

In view of this, and in a spirit similar to that of, inter alia, Eichengreen et al. (1985), and Dueker (and thus aggregate demand, thereby implementing desired stabilization policies) with smaller movements in the fed funds rate and other short term rates than would be required in a static framework (to induce a movement in the long term rates that is of the same magnitude).

¹²I make what is a standard assumption in the reaction function literature, namely that the policy instrument, inflation, and output gap variables are all stationary.

¹³The same is essentially the case for the averages of the effective fed funds rate, as Figure 2 demonstrates, and as is implied by the fact that the Fed's trading desk ensures that the actual fed funds rate never deviates too much from the set target.

(1999a,b), I propose a qualitative response approach that directly accounts for the discrete nature of the changes in the policy instrument of the Fed, and that explicitly distinguishes this variable, which is the one actually observed, from the Fed's *desired level* for the federal funds rate, ff_t^* , that is modeled as a latent variable, given by equation (2). While this latent variable is continuous, the observed policy instrument is discrete, and in particular it changes only when enough pressure for a change accumulates, that is, only when the difference between the actual value of the policy instrument and the level at which the Fed would like it to be is of a certain size. The size of that difference determines which, out of a small possible number of changes will take place.¹⁴

More specifically, the model I propose is a probit of a particular kind; namely it is multinomial and ordered because there are several possible outcomes (a modest number of possible amounts by which the Fed can decide to change its target for the fed funds rate), and these outcomes are ranked (as, for instance, it is recognized that an increase of 50 basis points is comparable to and ranks higher than an increase of 25 basis points).

In particular, we have that:

$$\Delta ff_t \in \text{category } j \text{ if } ff_t^* - ff_{t-1} \in (c_{j-1}, c_j), \quad j = 1, \dots, J, \quad (3)$$

where ff_t^* is the continuous latent variable of equation (2), and ff_t is *the observed policy instrument that changes only by one of J possible amounts at discrete points in time*, and where c_0, c_1, \dots, c_J are the threshold coefficients¹⁵ for movement between the J possible categories of change for the policy instrument. So, the difference $ff_t^* - ff_{t-1}$ represents the distance between the desired level of the fed funds rate this period (based on the reaction function of the Fed) and the actual level for the fed funds rate last period; thus, and as mentioned above, it can be seen as a measure of the "pressure" to change the fed funds rate in period t . The "intensity" of that pressure (that is, which of the (c_{j-1}, c_j) intervals (for $j = 1, \dots, J$) $ff_t^* - ff_{t-1}$ falls into) determines which of the J possible

¹⁴Accounting for this distinction, that is explaining why the Fed chooses to move its policy instrument rather infrequently, and only by one out of a small set of possible amounts when it does, rather than continuously, is beyond the scope of this paper. Possible models that could result in such Fed behavior include menu-cost type of models or models with the Fed as a strategic player using the target changes as signals of its policy stance.

¹⁵With $c_0 = -\infty$, and, $c_J = \infty$

changes will actually take place.

2.1 The Multiple Integral Problem

The obvious way to estimate the model of equations (2) and (3) is, as with all limited dependent variable models, maximum likelihood estimation. The likelihood (conditional on \tilde{X}_T , the explanatory variables for periods $1, \dots, T$) is the probability of the joint T -period event that has been observed, given \tilde{X}_T :

$$\Pr[\Delta f f_1 \in \text{category } j_1, \Delta f f_2 \in \text{category } j_2, \dots, \Delta f f_T \in \text{category } j_T \mid \tilde{X}_T].^{16}$$

Thus, this likelihood is a T -dimensional multiple integral with a T -variate Gaussian density as its integrand. The multiple integral problem arises in this context because, in contrast to standard linear autoregressive models, the Markov property does not hold here: The essential feature of linear time series that makes *conditional maximum likelihood estimation* (whereby the likelihood of the complete sample can be usefully expressed as a product of conditional likelihoods and of the density of the first observation¹⁷) easy is its Markov structure; that is, say for an AR(1), the period- t distribution conditional on period $t - 1$, does not depend on any other periods before period $t - 1$. This unfortunately is no longer true in our nonlinear context.

For example, let Ω_t denote the period t information set (that is defined to include the explanatory variables up to period t ,¹⁸ and the observed dependent variable up to period $t - 1$).

Then, the *conditional event probability for period t (given Ω_t)* is:

$$\begin{aligned} \Pr[\Delta f f_t \in \text{category } j \mid \Omega_t] &= \Pr[ff_t^* - ff_{t-1} \in (c_{j-1}, c_j) \mid \Omega_t] \\ &= \Pr[c_{j-1} + ff_{t-1} < ff_t^* < c_j + ff_{t-1} \mid \Omega_t] \quad , \quad j = 1, \dots, J \quad , \quad t = 1, \dots, T. \end{aligned} \tag{4}$$

¹⁶ j_1, \dots, j_T denote the categories in which $\Delta f f_1, \dots, \Delta f f_T$ actually fell in periods $1, \dots, T$.

¹⁷ See Hamilton (1994, Chapter 5) for a definition and discussion of conditional maximum likelihood of linear time series models.

¹⁸ Or, alternatively, it can be defined to include the explanatory variables for all periods up to T .

Solving equation (2)¹⁹ by backward substitution²⁰, we have that:

$$ff_t^* = \mu_t + \Gamma_t + \zeta_t ff_0^* + \xi_t, \quad (5)$$

where μ_t is a sum of constants, Γ_t is also a sum containing lags of the explanatory variables, ζ_t is a function of the coefficients of $\rho(L)$, ff_0^* is the dependent variable for the initial period, and ξ_t is a sum containing the error terms multiplied by the coefficients of $\rho(L)$. Now we can rewrite the period- t conditional event probabilities for $j = 1, \dots, J$, by plugging (5) into (4):

$$\begin{aligned} \Pr[\Delta ff_t \in \text{category } j \mid \Omega_t] & \quad (6) \\ &= \Pr[c_{j-1} + ff_{t-1} < \mu_t + \Gamma_t + \zeta_t ff_0^* + \xi_t < c_j + ff_{t-1} \mid \Omega_t] \\ &= \Pr[c_{j-1} + ff_{t-1} - \mu_t - \Gamma_t - \zeta_t ff_0^* < \xi_t < c_j + ff_{t-1} - \mu_t - \Gamma_t - \zeta_t ff_0^* \mid \Omega_t]. \end{aligned}$$

Inspecting equation (6), we can gain some more intuition on the *multiple integral problem*, that arises in this context and that greatly complicates the estimation task.

While ξ_t does satisfy the Markov property, Δff_t does not. As we can see from equation (6) the conditional period- t event probability that $\Delta ff_t \in \text{category } j$ explicitly depends not only on Δff_{t-1} , but on information from all previous periods as well. Indeed, and while we need only the *conditional* distribution of ξ_t given Ω_t for the probability of equation (6), the Markov structure of ξ_t does not help much with this problem. The information set Ω_t includes values of the observed dependent variable up to period $t - 1$. However, Δff_{t-1} places bounds only on ξ_{t-1} (because of equation (3)), but does not pin it down completely, and thus all the past history of the dependent variable contains additional useful information regarding ξ_{t-1} . That is, the t -period event probability is still a t -dimensional multiple integral *even when this probability is conditioned on Ω_t* . And, as the error terms are Gaussian, these multiple integrals have multinomial Gaussian pdf's as their integrands,

¹⁹For expositional simplicity I consider the case of Equation (2) with one lag only.

²⁰Writing the latent equation in levels is a useful expositional device that Eichengreen et al (1985) adopt to illustrate the multiple integral problem. Here I adapt their approach to the context of this study.

and thus there is no closed-form solution for such multiple integrals.

To further see the intuition behind this, let us consider an example: Let's say the Fed raised the policy instrument in period $t-1$. While f_{t-1}^* is not explicitly revealed by this action, it is likely to be at the lower end of its possible values implied by the new category in which it now falls (equation (3)). This, of course, reduces the probability of a further increase of the policy instrument in period t , and must thus be taken into account when assessing the probability of a further increase in period t . Thus, it is essentially the entire history (all the periods) that must be taken into account when assessing these probabilities, precisely because the lagged latent variable is not observable.

Thus, the question that arises is how to proceed with estimation when faced with a likelihood that contains multiple integrals of such high-dimensionalities.

Standard numerical integration techniques, such as Gaussian quadrature, can effectively handle only up to four or five multiple integrals, and thus are essentially not feasible for multiple integral dimensionalities such as the ones of the present context. An early notable attempt to tackle the multiple integral problem is Eichengreen et al. (1985) who however introduce discrete approximations of the multiple integrals and face a rather steep trade-off in terms of accuracy versus computational cost.

Classical and Bayesian simulation techniques that have been developed recently greatly dominate such numerical approaches (in terms of their computational costs, accuracy, and in terms of their potential to handle modeling of complicated dynamics) and have thus become by far the dominant approach in recent literature.

One influential approach is that of the simulation-assisted extremum estimation techniques such as the Methods of *Simulated Likelihood, Moments, or Scores* that have been developed in the microeconometrics literature (see, inter alia, Lerman and Manski (1981), McFadden (1989), Pakes and Pollard (1989), Hajivassiliou and McFadden (1998)). Simulators, such as the *frequency simulator* of Lerman and Manski (1981), or the *GHK simulator* developed by Geweke (1991), Hajivassiliou (1990) and Keane (1994) are used to estimate the multiple integrals, and thus to obtain simu-

lated objective functions²¹, which, in conjunction with numerical optimization techniques, provide a framework that can deliver estimates of the parameters of interest. However, these methods have been employed so far mostly in micro contexts²², where the dimensionality of the multiple integrals is typically smaller than in time series contexts. Indeed, the computational costs associated with the required simulations, combined with similar costs of the needed numerical optimization techniques can render them impractical, or even infeasible in time series contexts such as the one of this paper.

Thus, the approach that this paper proposes is an MCMC approach that overcomes the multiple integral problem through a *Gibbs sampling algorithm with data augmentation*, whereby simulated samples of the latent variables $ff_t^*, t = 1, \dots, T$, are generated through their model-implied conditional distributions. This approach also avoids the additional complications associated with numerical optimization, as it falls outside the extremum framework. Finally, an important additional advantage of the MCMC approach is its flexibility: The block structure of MCMC implies that additional time series features such as Time Varying Parameters, or Regime Switching can be readily incorporated into the benchmark model, whereas attempting something similar in an extremum framework is likely to be substantially more cumbersome. The specifics of the MCMC approach used in this paper are described next.

2.2 The MCMC Estimation Strategy

Let θ denote the parameter vector, that in the context of the benchmark specification is taken to include the coefficients of the explanatory variables, the variance, and the latent variables. Let

²¹The frequency simulator exploits the fact that the event probabilities can be viewed as expectations of indicator functions, and thus employs accept-reject types of algorithms to construct estimates of such expectations. Specifically, and in the present context, we have that the likelihood can be written as follows:

$\Pr[\Delta f f_t \in \text{cat } .j_t, t = 1, \dots, T \mid \tilde{X}_T] = \int_{L_T}^{U_T} \int_{L_{T-1}}^{U_{T-1}} \dots \int_{L_1}^{U_1} f_T(\xi_1, \dots, \xi_T) d\xi_1 \dots d\xi_T = E(I(\xi))$, where $I(\xi)$ is an indicator function that is 1 if T -dimensional ξ falls within the T -period limits of integration, and 0 otherwise. So, $E(I(\xi))$ can be

estimated by the frequency simulator : $\frac{\# \text{ accepts}}{\# \text{ draws}} = \frac{1}{R} \sum_{r=1}^R I(\xi_r)$. However, and as is evident from the above expression, the frequency simulator is computationally inefficient, especially for low event probabilities, and it is a jump-function of the parameters. It is dominated by other simulators that have been developed recently, and in particular by the GHK simulator, which provides a feasible framework for simulating multiple integrals by transforming the task of sampling from multivariate truncated distributions (that is required to simulate the multiple integrals) to one of sequentially sampling from conditional univariate truncated distributions (that can be easily performed through the inversion of the cdf's of these truncated univariate distributions).

²²One of the very few exceptions is Lee (1999).

$p_T(\theta)$ denote the posterior distribution of the parameter vector. MCMC techniques are in general designed to generate simulated samples that are approximately distributed from $p_T(\theta)$, *without directly drawing these samples from $p_T(\theta)$ per se*. The specific MCMC technique used in this paper is the *Gibbs sampler*: It is particularly suitable in contexts such as the present one, where evaluating the entire joint posterior distribution of the parameters at any given point of the support space may be very hard (in the present context, that posterior distribution is proportional to the likelihood, which, as analyzed above, contains high-dimensional multiple integrals), but where it is easy to sample from marginal and joint conditional posterior distributions of *subsets (or blocks) of the parameter vector*, because of simplifications that occur in these conditional posteriors.

The simulated sample that is produced by the Gibbs sampler constitutes an ergodic Markov Chain, whose stationary distribution is $p_T(\theta)$. In other words, it is a Markov Chain that converges to an invariant distribution ($p_T(\theta)$) independently of the initial conditions used to initiate the chain. Thus, and after a "large enough" number of simulations, that is, after a certain number of pre-convergence "burn-in" draws have been discarded, a sample of simulated values can be created, and the joint and marginal posterior distributions of the parameters can be approximated by the respective joint and marginal empirical distributions of these simulated values.

The ergodicity property is central as it ensures that the initial conditions which are typically arbitrary have no effect on the results, and it also serves as a useful tool to assess whether convergence has been achieved or not: If the results change when the Gibbs sampler is initiated from different starting points, then this typically implies that convergence has not been achieved. MCMC convergence diagnostics is in general an active research area, and various approaches have been proposed in the literature²³. The approach adopted here is similar to that of McCulloch and Rossi (1994), whereby the empirical distributions of the simulated values are compared when the Gibbs sampler is initiated from different starting points, and as the number of simulations increases, looking for evidence of non-trivial changes in these distributions.

Once convergence has been established according to the criteria just outlined, statistics based

²³Robert and Casella (1999), Chib (2001), and Geweke and Keane (2001) are some of the excellent sources for more information on MCMC convergence diagnostics, as well as on MCMC theory.

on the simulated sample can be constructed and can serve as estimators of the parameters of interest. An important additional computational advantage of this approach is that it falls outside the extremum context, and thus no costly numerical optimizations are required to obtain the estimates. In the following sections, the estimates of the parameters of interest are the means of their simulated marginal posterior distributions, and confidence intervals are constructed using the quantiles of these posterior distributions²⁴.

Note that, while MCMC techniques in general, and the Gibbs sampler in particular are encountered almost always in Bayesian contexts in the literature, this paper adopts a *classical perspective*. The estimator just described is a Bayes estimator. Under suitable regularity conditions and as the sample size increases, the likelihood overwhelms the priors, whose effect thus fades away, and thus the Bayes estimator, *viewed from a sampling perspective*, is asymptotically efficient. The intuition for this is very much in the tradition of Laplace, who observed that posterior distributions start to look increasingly like Gaussian distributions as the sample size grows. Bernstein (1917) and von Mises (1931) are early contributions that formalized Laplace’s intuition. More recent contributions, such as Lehmann and Casella (1998), derive the asymptotic efficiency of Bayes estimators under an appropriate set of conditions. Chernozhukov and Hong (2003) generalize past results of the literature as they propose a class of estimators (which they call *Laplace type* estimators) which use general statistical criterion functions that include parametric likelihood functions (that past literature has mostly focused on) as special cases. Monokroussos (2004) specializes the theory to the particular class of Limited Dependent Variable models of Time Series (that includes the ordered probit of this study as a special case), and also examines the small sample biases associated with such estimators. Stated succinctly, and for the present context, the proposed estimator provides a feasible way to estimate probits of time series, and is also asymptotically equivalent to the maximum likelihood estimator.

Previous work that employs Gibbs sampling techniques for Limited Dependent Variable and Time Series models includes Albert and Chib (1993a,b) and Dueker (1999b). The specifics of the

²⁴As Chernozhukov and Hong (2003) and Monokroussos (2004) discuss, the asymptotic validity of such confidence intervals for the present context is based on the fact that the Information Equality holds in this context.

Gibbs sampling algorithm that is used here for the benchmark model are as follows.

Divide the set of parameters into a multi-block setup of one block per latent variable, $\lambda_{1_t} = \{ff_t^*\}$, $t = 1, \dots, T$, one block for the variance, $\lambda_2 = \{\sigma^2\}$, and one block for the coefficients of the explanatory variables, $\lambda_3 = \{\alpha, \beta, \gamma, \rho_1, \dots, \rho_n\}$. Then:

Step 1: Specify arbitrary initial values $\lambda_1^{(0)}, \lambda_2^{(0)}, \lambda_3^{(0)}$, and set $i = 0$.

Step 2: Cycle through the following conditional distributions, drawing:

$$\begin{aligned}
 \lambda_{1_1}^{(i+1)} & \text{ from } p_{ff_1^*}(\lambda_{1_1} | \lambda_{1_2}^{(i)}, \dots, \lambda_{1_T}^{(i)}, \lambda_2^{(i)}, \lambda_3^{(i)}, Y_T) \\
 \lambda_{1_2}^{(i+1)} & \text{ from } p_{ff_2^*}(\lambda_{1_2} | \lambda_{1_1}^{(i)}, \lambda_{1_3}^{(i)}, \dots, \lambda_{1_T}^{(i)}, \lambda_2^{(i)}, \lambda_3^{(i)}, Y_T) \\
 & \vdots \\
 \lambda_{1_T}^{(i+1)} & \text{ from } p_{ff_T^*}(\lambda_{1_T} | \lambda_{1_1}^{(i)}, \lambda_{1_2}^{(i)}, \dots, \lambda_{1_{T-1}}^{(i)}, \lambda_2^{(i)}, \lambda_3^{(i)}, Y_T) \\
 \lambda_2^{(i+1)} & \text{ from } p_2(\lambda_2 | \lambda_{1, \dots, T}^{(i+1)}, \lambda_3^{(i)}, Y_T) \\
 \lambda_3^{(i+1)} & \text{ from } p_3(\lambda_3 | \lambda_{1, \dots, T}^{(i+1)}, \lambda_2^{(i+1)}, Y_T),
 \end{aligned} \tag{7}$$

where Y_T denotes the entire history of the data for periods $1, \dots, T$, and superscript i indicates the iteration of the Gibbs sampler. This choice of blocks is dictated by the fact that the resulting conditional posteriors are easy to sample from, and the Appendix provides the distributional details of these posteriors, and also other details of the Gibbs sampler, including robustness checks performed using different priors.

The multi-block setup of one block per latent variable is employed to implement the technique of *data augmentation*, introduced by Tanner and Wong (1987), whereby the latent variables are generated from their model implied conditional distributions $p_{ff_1^*}, \dots, p_{ff_T^*}$. Note that these distributions are conditioned on the entire history of the data and thus that a smoothing algorithm is required. The standard way to approach a problem of this sort is with a state-space framework and using the Kalman filter; however, it is unclear how to usefully cast the system in a state-space form

so as to employ the Kalman filter in this context²⁵. I thus implement the smoothing algorithm for the latent variables by exploiting simplifications that occur in the conditional distributions of these latent variables, and then observing that functional forms for these simplified conditional distributions can be obtained from the joint distribution of all the error terms where each latent variable appears. This smoothing algorithm for the latent variables is derived in the Appendix.

3 Data

The dependent variable is the Fed's policy instrument and the explanatory variables are, in addition to lags of the Fed's policy instrument, real-time forecasts of inflation and of the output gap. The data series used span the period from January of 1969 to June of 1998²⁶. The rest of this section provides details on these time series:

The Fed's Policy Instrument

The Fed's policy instrument is typically assumed to be the federal funds rate. This is the working assumption of most of the applied work on reaction functions and is in line with the prescriptions of past literature (such as Goodfriend (1991), Bernanke and Blinder (1992), Bernanke and Mihov (1998)). However, and while the literature has mostly focused on choosing a variable that is a sufficient indicator of monetary policy and has largely settled on the federal funds rate as being the best choice, it has generally neglected²⁷ the issue of characterizing the specifics of its time series behavior:

The Fed announces periodically (usually during FOMC meetings, but occasionally in between FOMC meetings too) a target for the federal funds rate. These announcements are eagerly anticipated and closely scrutinized by the financial markets internationally, precisely because it is these announcements and the associated movements in the target that are understood to be the best,

²⁵Note that the obvious choice, namely, for the transition equation to be the latent equation (2) and for the measurement equation to be equation (3) is not feasible because equation (3) is not linear.

²⁶Greenbook forecasts (which are used for the construction of the explanatory variables) are made available to the public with a 5 year lag.

²⁷With only a few exceptions, such as Hamilton and Jordà (2002), and Dueker (1999a).

most direct indicators of the U.S. monetary authority's policy stance. As argued earlier, however, the target moves at discrete points in time only, and its movements fall into one out of a small possible number of specific categories of change. This pattern is largely preserved even when periodic averages of the fed funds rate are considered, as it is the task of the trading desk of the New York Fed to ensure that the fed funds rate never deviates much from the target at any point in time.

Thus, an ordered probit reaction function such as the one given by equations (2) and (3), estimated using the time series of the changes of the policy instrument as the dependent variable, is better suited to capture the discreteness of the policy instrument.

However, the time series on the target and on its changes are not available for the entire period considered in this study, and not for the early part of the sample in particular. The Fed began to announce the target explicitly only in 1994, although, even before 1994 changes in the target were quite accurately inferred by the market and indeed news on such changes would even be published in the financial press. Researchers have used such historical information to compile time series of target changes for periods prior to 1994 (such as Cook and Hahn (1989) for the mid and late 1970's and Rudebusch (1995) for the mid-1980's and later), but there is no complete such series for the entire period considered in this study.

Thus, the dependent variable that I use is the (annualized) averages of the fed funds rate at the "FOMC frequency", that is, monthly averages for the months during which there was an FOMC meeting²⁸. These averages, for which there is a complete series for the period of interest, can thus be viewed as proxying for averages of the target. One would expect, given that the Fed ensures that the fed funds rate never deviates too much from the set target at any given period, that the loss from such a strategy is only minimal. Indeed, and as Figure 2 (which maps monthly averages of the fed funds rate against monthly averages of the target for the Greenspan period) illustrates, the fed funds rate tracks the target quite closely²⁹. Additional evidence to this is provided in Table

²⁸The choice of frequency is dictated by the explanatory variables: In particular the Greenbook forecasts (from which the explanatory variables are constructed) are available only for the months during which there was an FOMC meeting.

²⁹This can essentially be considered to be a case of measurement error in the dependent variable. Such error in the dependent variable in the ordered probit context may not be as innocuous as it would be in a linear context though because of the possibility that the measurement error "pushes" the discrete dependent variable into a different

1, which contains estimates of the benchmark reaction function for the Greenspan period using the monthly averages for the fed funds rate; these estimates are quite similar to their respective estimates obtained when using the target as the dependent variable, either as averages at the FOMC frequency, or as end-of-month values for months during which there was an FOMC meeting.

Real-time data on Fed forecasts of inflation and of the output gap

The data used are real-time Greenbook forecasts of the Federal Reserve for the GDP deflator and for unemployment. The Greenbooks and the forecasts contained therein are prepared by staff of the Federal Reserve before each Federal Open Markets Committee meeting, and thus contain estimates based on real-time information. The forecasts are for varying forecast horizons, but only short horizons are consistently available - in this study I use three-quarter-ahead forecasts for the benchmark specification, but I also conduct robustness checks using data for zero, one and two quarters ahead.

The inflation variable used is based on the GDP deflator (up to October of 1991) and the GNP deflator after that date³⁰. Also, and following past literature, such as Boivin (2004), and Orphanides (2002), I construct a proxy for the output gap variable using forecasts for unemployment, and specifically, I take the output gap at time t to be proxied by the difference between the natural unemployment rate (the time t natural rate of unemployment is taken to be the historical average of the unemployment series up to that point) and the forecast for the unemployment rate³¹ at time t .³²

category than the one it would otherwise fall into. However, and conversely, if this change of categories never happens, or happens only rarely, as is the case with the Greenspan sample I considered, then there will be little cost associated with such measurement error.

³⁰This is dictated by the availability of these variables, as GNP deflator data replace GDP deflator data in the Greenbooks starting October, 1991, and it is likely to have only a negligible effect on the results.

³¹This is standard practice in the literature and indeed Okun's Law guarantees that there is a close relationship between this "unemployment gap" proxy and the output gap.

³²I also conducted robustness checks using the alternative measure of real output gap of Orphanides (2004): There are two problems with using the Orphanides data set in the present context, however: First, the quarterly frequency of that data set, as opposed to the "FOMC frequency" of the data set used in this study, is likely to confound variations in the changes of the policy instrument by bundling together successive and distinct such changes. Second, a complete series is available only starting with the 3rd quarter of 1973. This is likely to further compromise the validity of the results in the pre-Volcker period (which ends in 1979) as the 1973-1979 sample, (with a quarterly frequency), may be too short to identify the slope coefficients of the reaction function. The results using the Orphanides data set are quite similar to the results of the rest of this study for the Volcker-Greenspan period. For the pre-Volcker period, the

The real-time nature of these forecasts is in line with prescriptions of recent literature (see, inter alia, Orphanides (2001, 2002, 2004)) according to which reaction functions ought to be estimated using data that were available to Fed policy makers at the time their decisions were made, if a meaningful historical analysis of policy is to be made. Estimating these functions with revised data, which were only available typically much later than in real time, can seriously confound important aspects of the policy making process, especially when, as is often the case, the revised series differ substantially from their real-time counterparts.

The use of forecasts results in a forward-looking reaction function, which is arguably a more realistic specification in the face of lags that are generally thought to be present in the monetary transmission mechanism, and it also encompasses backward looking specifications, in the sense that it is likely that the Fed, when constructing its forecasts, has a broad information set that includes the variables used in a backward looking specification.

However, this approach is not free of pitfalls. In particular, there isn't much detailed information in the public domain regarding how these Greenbook forecasts are constructed, and it is likely that quantitative methods are employed, and that there is also a strong judgmental component. Also, and as Boivin (2004) argues, it is likely that most of these are conditional forecasts, with one of the conditioning assumptions being that the federal funds rate will remain unchanged over the horizon that the forecast spans. This is of course a restrictive assumption as it is unlikely that the FOMC considers forecasts based only on this particular scenario (that the fed funds rate will not change in the near future), and thus it may compromise the validity of estimates of forward-looking reaction functions that use such Greenbook forecasts.

Thus, and while this may be an imperfect approach, it is not clear that a popular alternative that has been proposed in the literature (and which is the approach that CGG adopt) is preferable: That alternative approach treats the expectations of a standard forward looking reaction function, such as equation (2') below:

results obtained using the Orphanides quarterly data for the 1973-1979 period differ from the results of this study for the 1969-1979 period (at the FOMC frequency).

$$ff_t^* = \alpha + \rho(L)ff_{t-1}^* + \beta E[\pi_{t,h}|\Omega_t] + \gamma E[y_{t,h}^{gap}|\Omega_t], \quad (2')$$

as unobservable, and thus replaces them with the realizations of the inflation and output gap variables, and it relies on a rational expectations assumption to derive a set of orthogonality conditions between the resulting error term and Ω_t , the information set of the Fed at time t . These orthogonality conditions can then be exploited for estimation using a Generalized Method of Moments approach, with instruments drawn from the Fed's information set. However, recent work (such as Mavroeidis (2002, 2004)) has shown that this strategy is plagued by weak identification problems, associated with the unrealistic assumption of lack of feedback (which of course is a problem with the Greenbook forecast approach too), and also with the instruments actually used for the GMM estimation being weak.

4 Estimation Results and Discussion

The benchmark model is equations (2) and (3). It is assumed that there are 5 categories of change for the observed variable, namely a no change category, a small change (positive or negative) category, and a bigger change (positive or negative) category. While it might be possible to create an algorithm that endogenizes the number of categories, there seems to be little loss from this assumption of 5 categories of change in the present context: One realizes, simply by observing the actual time series of target changes that there have been no changes in the target³³ for roughly half of the time, and for the other half of the time there have been either small changes in the target, or bigger changes, that is, changes whose magnitude has been greater than one times 25 basis points after 1989, and greater than three times 6.25 basis points before 1989.

Thus the benchmark specification is one with these 5 categories of change, and with the threshold coefficients being -18 , -5 , 5 , and 18 basis points. Thus, and for example, the "no-change" category is the one where the change in the observed variable is between -5 and 5 basis points. While

³³Or almost no changes when FOMC-monthly averages of the fed funds rate are used.

there might be little loss associated with exogenously imposing a number of categories of change in the present context, the same is not necessarily true with exogenously imposing the threshold values that define these categories. Thus, the following section reports results from performing robustness checks on the threshold values and also on attempts to estimate such threshold values (by augmenting the Gibbs sampler for the benchmark model with additional blocks from which the threshold coefficients are generated).

The forecast horizon that the benchmark specification assumes is $h = 3$ quarters ahead. This is consistent with current practices in the forward-looking reaction function literature³⁴, but I also consider alternative specifications with different forecast horizons in the following section.

The results for the benchmark model are reported in Tables 2 and 3. These tables contain panels with quantiles and statistics of the posterior distributions of the parameters of interest corresponding to different numbers of iterations of the Gibbs sampler³⁵. Several panels such as those of Table 2, together with their respective histograms of the posterior distributions, such as those Figures 3A and 3B³⁶, are produced for different starting values that initiate the Gibbs sampler and for different numbers of iterations of the sampler, and they are compared and evidence of any substantial changes in the results is sought for. For all the results reported in this study, the results change very little as different starting values and different numbers of iterations are considered. This serves as evidence of convergence, at least according to the criteria outlined in the previous section.

Table 2 provides the results for the benchmark model and for one lag of the latent dependent variable, and Table 3 provides the results for the benchmark model with 2 lags (with the second lag being insignificant³⁷). The results from both of these tables are similar and they point towards

³⁴See Boivin (2004) and Orphanides (2004).

³⁵The results reported here are based on 1300 iterations, with 300 "burn-in" states, as all the available evidence in various experiments I conducted showed that this was a sufficient number of iterations (based on the criteria that are used in this study to established convergence (that were outlined earlier), and also based on experiments and comparisons with higher numbers of iterations).

³⁶The figures are based on all the iterations and provide a more complete picture of the marginal distributions. These figures are plain histograms; no nonparametric techniques for selecting the bandwidth have been employed. Our main focus when examining these figures is to look for evidence of more than one modes; thus our goal is to preserve the rough shape of the density, minimizing biases; the side effect of that is that our graphs are too wiggly; thus, these wiggles and spikes are just a result of the way the histograms are generated.

³⁷The two-lag specification is the one used by Clarida, Galí and Gertler (2000)

substantial differences in monetary policy between the two periods before and after Paul Volcker's appointment as chairman of the Fed in August of 1979:

Specifically, the pre-Volcker period (which, for the sample considered in this study covers the last two years of Martin's tenure, Burns's tenure in the 1970's, and Miller's tenure as Chairman of the Fed from March of 1978 to August of 1979) is characterized by a weak overall response to inflation: For instance, for the specification with one lag, the inflation coefficient is not significantly different from 0, and the implied long-run inflation coefficient is $\frac{\beta}{1-\rho_1} = \frac{0.0682}{1-0.8664} = 0.51048$. This is well below 1 and thus does not satisfy the Taylor property; it suggests that the pre-Volcker Fed "accommodated" inflation, as it would raise the nominal interest rate by less than increases in expected inflation, thus effectively allowing the real short-term interest rate to decline.

These results contrast starkly, however, with the estimated response of the Fed towards inflation during the Volcker-Greenspan era. For the 1 lag specification again, the inflation coefficient is substantially higher than before and is now significant, and the implied long-run inflation coefficient is now $\frac{0.4518}{1-0.7679} = 1.9466$, which suggests that the Volcker-Greenspan Fed adopted a very strong anti-inflationary stance, as it would substantially raise *both* the nominal *and* the real interest rate in response to increases in anticipated inflation.

Quite clearly, these findings confirm the central message of CGG that there were substantial differences in the Fed's response towards expected inflation between the two periods. These results are thus compatible with CGG's theory that the pre-Volcker's Fed accommodative policy stance was destabilizing, as it allowed for the possibility of expectations-based inflationary spirals, and also as it was less effective in countering negative shocks to the economy, and that the dramatic shift towards a much stronger anti-inflationary stance of the Fed that took place when Paul Volcker became chairman of the Fed was a central factor behind the improved macroeconomic outcomes and the greater stability of the 1980's and 1990's.

Conversely, and as can be seen in Tables 2 and 3, the coefficient of the output gap was low (0.1144 in the 1-lag specification of Table 2) and insignificant during the Volcker-Greenspan period, and substantially higher (0.3006 in the 1-lag specification) and significant during the pre-Volcker

period. Thus, and on the basis of these results, it can be argued that the Fed effectively pursued a pure inflation targeting policy during the 1980's and 1990's, while it was much more responsive towards the output gap during the pre-Volcker period.

Given that the data used for the construction of the output gap variable are real-time Greenbook forecasts on unemployment, and since, as demonstrated by Orphanides (2002), estimates of the output gap based on Greenbook data on unemployment are upwardly biased, the results of Tables 2 and 3 on the coefficient of the output gap are consistent with the explanation put forth by Orphanides in his recent papers (Orphanides (2002, 2004)). Orphanides argues that the main difference between the two periods before and after Volcker's appointment is that the pre-Volcker Fed was essentially too activist in its response to real-time output gap estimates that ex-post proved to be overambitious, and thus that this excessive activism of the pre-Volcker Fed contributed to the poor macroeconomic outcomes and the Great Inflation of the 1970's. Following Volcker's appointment the Fed adopted a more cautious approach towards possibly inaccurate real-time estimates of the output gap and in general a more realistic stance in the sense that it better recognized the limitations of monetary policy in attempting to achieve output stabilization. However, and in contrast to CGG, Orphanides finds that the Fed had a strong anti-inflationary stance both in the 1960's and 1970's and later.

Both CGG and Orphanides essentially adopt a "policy mistakes"³⁸ view to explain the Great Inflation and the contrast in macroeconomic performance between the 1970's and the 1980's-1990's in that they argue that the pre-Volcker Fed made mistakes in its conduct of monetary policy, and that starting with Volcker's appointment the Fed to a large extent avoided mistaken practices of the past. CGG and Orphanides differ, however, in terms of their assessments on what these policy mistakes were.

The results presented here provide further evidence for such a policy mistakes view in terms of *both* the inflation *and* the output gap variables. They are thus compatible with more tightly parameterized approaches, such as learning models in which the monetary authority initially has

³⁸The terms "policy mistakes view" and "bad luck view" that follows are taken from Primiceri (2004).

wrong perceptions in real time about both the output gap and the output-inflation trade-off and eventually corrects such misperceptions.

The results of the benchmark model are not necessarily inconsistent however with alternative explanations such as the "bad luck" view that focuses on the role of unfavorable non-policy shocks in the 1970's, or a more gradualist approach that would argue that any policy changes were more gradual than the stark contrast described by CGG. Thus the section that follows takes up the task of investigating to what extent the main results of the benchmark model stand when extensions of this sort and other robustness checks are considered.

5 Extensions and Robustness Checks

Various robustness checks and extensions are considered in this section, and the main results and conclusions obtained from the estimation of the benchmark model of the previous section are mostly preserved in all of these exercises. The results that follow are obtained from estimates of the benchmark model using different forecast horizons, different threshold values that define the categories of change of the observed dependent variable, and also, from an extension of the benchmark model that allows for modeling of possible heteroskedasticity, and an additional extension that allows for time-varying parameters. The rest of this section considers each of these robustness checks and extensions in turn:

5.1 *Different forecast horizons*

The forward-looking specification for the latent equation of the benchmark model (equation (2)) is too restrictive in that it assumes that the Fed's responses to forecasts of inflation and of the output gap are independent of the forecast horizon. This is clearly an unrealistic assumption as the Fed realizes that a change in its policy instrument at any given point in time will have different effects on the macroeconomy at different points in the future. Indeed, a more realistic depiction of the Fed's forward-looking behavior results if we replace equation (2) by the following:

$$ff_t^* = \alpha + \rho(L)ff_{t-1}^* + \beta_h\pi_{t,h} + \gamma_h y_{t,h}^{gap} + \varepsilon_t^*, \quad (2'')$$

with β_h and γ_h varying with the forecast horizon h .

Similarly, and given that actual Greenbook forecasts are used in the estimation, and since, as discussed earlier, Greenbook forecasts are often conditioned on particular paths of the policy instrument (such as a path of no change for the policy instrument), and that the monetary policy makers are unlikely to consider single-scenario conditional forecasts when reaching their policy decisions, it is likely that this is a misspecified framework, with the effects of such misspecifications varying with the forecast horizon.

Thus, and as the choice of the forecast horizon for the benchmark model was arbitrary, it is essential to consider how the results obtained using the benchmark specification (that assumes a forecast horizon of $h = 3$ quarters ahead) compare to those of alternative specifications with different forecast horizons. I conducted experiments using Greenbook forecasts on $h = 0, 1, 2$ quarters ahead³⁹ and I report the results in Table 4. The results are similar to those obtained with the 3-quarter ahead forecast horizon (although the inflation coefficient tends to be somewhat more significant with these alternative specifications than with the benchmark specification during the pre-Volcker era) and all the conclusions from above remain unchanged.

5.2 *Threshold coefficients and identification*

The estimation of the benchmark model is done with the threshold coefficients being set to arbitrary values, rather than being estimated, and the obvious concern here is that the results may be affected by the choice of values for the threshold coefficients. It is in principle possible to estimate the threshold coefficients of equation (3) by augmenting the Gibbs sampler for the benchmark model with appropriate additional blocks, and the Appendix illustrates how this is done. However, such estimation attempts are plagued by identification problems.

³⁹The Greenbooks occasionally provide forecasts for longer forecast horizons too. However, and starting with 4 quarters ahead, there are too many gaps in these time series.

First of all, a setup where all the threshold coefficients, together with all the other parameters, are all being freely estimated, is not identified, as we have an intercept in our model: Adding a constant to all three sides of the double inequality giving the conditional event probability (equation (4)) will not change that event probability, and thus the likelihood will not change either. Indeed, a simple inspection of the histograms of the posterior distributions associated with such estimation attempts strongly suggest lack of identification as there is clear evidence of lack of convergence of the Gibbs sampler and of multimodality in the conditional posterior distributions for the threshold coefficients.

Fixing one of the threshold coefficients addresses the identification issue in a strict sense, but the evidence I get from performing the usual convergence diagnostics checks and from examining the histograms of the posterior distributions is that nevertheless there isn't improvement from the previous case. This suggests that weak identification of the threshold coefficients is an important factor in our context. This sounds quite intuitive too: For instance, altering the threshold coefficients mildly is unlikely to change the event probabilities, and thus the likelihood, substantially.

Thus there seems to be little loss associated with fixing the threshold coefficients at some reasonable values. As mentioned earlier, the values used in the benchmark specification are: -18 , -5 , 5 , and 18 basis points. There is clearly very little loss associated with -5 , 5 : These numbers are wide enough so as to capture the no-change category when the series used is the monthly averages of the effective fed funds rate and they are also too small for the small-change categories (any changes in the target have historically come in multiples of 6.25 or 25 basis points). However, the same may not be true for the numbers chosen to divide the small change and bigger change categories in the benchmark specification, namely -18 , 18 ; for instance, there have been several target changes that took place before 1989 and that were equal to ± 18.75 basis points. Thus I examine the robustness of the results when these numbers are altered by re-estimating the benchmark specification using ± 25 instead of ± 18 basis points to divide the categories of small change and of bigger change. The results from this experiment are reported in Table 5 and are very similar to those of the benchmark specification.

5.3 Modeling heteroskedasticity

An influential alternative, and to some extent, competing view that has been developed recently in the literature⁴⁰ argues that it was exogenous, non-policy shocks, whose volatility was higher in the 1970's than in the 1980's and 1990's, that were the main cause of the Great Inflation and of the contrast in macroeconomic performance between the pre-Volcker and the Volcker-Greenspan periods. In view of this, it is important to test the robustness of the results obtained with the benchmark model when this is embedded in a richer framework, that allows for possible heteroskedasticity.

I allow for such potential heteroskedasticity by using a two state regime switching framework for the variances⁴¹. Specifically, the model that I estimate is the same as the benchmark model with the latent equation now being as follows:

$$ff_t^* = \alpha + \rho(L)ff_{t-1}^* + \beta\pi_{t,h} + \gamma y_{t,h}^{gap} + \varepsilon_t^*, \quad (2)$$

where

$$\varepsilon_t^* \sim N(0, \sigma_{S_t}^2),$$

and where

$$\sigma_{S_t}^2 = \sigma_0^2(1 - S_t) + \sigma_1^2 S_t = \sigma_0^2(1 + h_1 S_t), \quad \text{with } \sigma_1^2 = \sigma_0^2(1 + h_1),$$

where S_t is a hidden Bernoulli random variable with the following transition probabilities:

$$\Pr[S_t = 0 | S_{t-1} = 0] = q, \quad \Pr[S_t = 1 | S_{t-1} = 1] = p.$$

I also assume that $p, q \in (0, 1)$, which ensures that neither state is transient⁴². Furthermore I assume that $h_1 > 0$. This is a convenient normalization condition that deals with the "label

⁴⁰See, inter alia, Kim and Nelson (1999b), Sims (1999), Blanchard and Simon (2001), Stock and Watson (2003), and Sims and Zha (2004).

⁴¹See Hamilton (1989, 1994).

⁴²This assumption is required for identification, since if $S_t = 0$, for all t , or if $S_t = 1$ for all t , then the two separate variances are not identified.

switching" issue that is present in mixture models⁴³.

The Gibbs sampler that is used for this model is that of the benchmark model that is augmented to include the following blocks⁴⁴:

$$\lambda_{2'}, \lambda_{2''} = \{\sigma_0^2, \sigma_1^2\}$$

$$\lambda_{4_{1,\dots,T}} = \{S_t\}, t = 1, \dots, T$$

$$\lambda_{5'}, \lambda_{5''} = \{p, q\}$$

The smoothing algorithm that is used for the generation of the state variables is based on a combination of a multi-move Gibbs sampling step and Hamilton's (Hamilton (1989)) filter. The specifics of this algorithm and of the rest of the Gibbs sampler for the regime switching model are provided in the Appendix.

The results are provided in Table 6, and they are quite similar to those of the benchmark model. The main finding of a stark, dramatic contrast in US monetary policy between the pre-Volcker and the Volcker-Greenspan periods receives at least as much support as with the benchmark model. Interestingly, and as the two figures associated with Table 6 illustrate, we are in the low-volatility state for the majority of the time in the pre-Volcker period, and we are in the low-volatility state almost all the time in the Volcker-Greenspan period.

While it would be useful to consider specifications that allow for richer dynamics too, such as, for instance, a regime framework for the second moments with a higher number of states, the results obtained with this model serve at least as an indication that heteroskedasticity is not a major factor in this context.

⁴³See Hamilton et al. (2003) for further details.

⁴⁴The $\lambda_{2'}, \lambda_{2''}$ blocks replace λ_2 .

5.4 Time Varying Parameters

An interesting extension to the standard reaction function literature has been proposed by, inter alia, Cogley and Sargent (2001, 2002), Boivin (2004), and Jalil (2004). According to this approach, there may have been richer dynamics in US monetary policy than what can be captured by a simple split-sample strategy such as that of CGG, and also, policy changes may have been slower and more gradual than the split-sample approach can uncover. Thus, these authors, motivated by such considerations, estimate reaction functions with Time Varying Parameters.

As the benchmark model of this study is estimated within a split sample framework, it would be interesting to see if the results and conclusions from the estimation of the benchmark model are preserved when Time Varying Parameters are considered. Thus, the TVP version of the benchmark model is equation (3) and a TVP version of equation (2):

$$ff_t^* = \alpha_t + \rho_t(L)ff_{t-1}^* + \beta_t\pi_{t,h} + \gamma_t y_{t,h}^{gap} + \varepsilon_t^*. \quad (2''')$$

Following the authors mentioned above, I model the Time Varying Parameters as driftless random walks.

So, let Φ_t be a $(n+3) \times 1$ vector: $\Phi_t = [\alpha_t \ \rho_{1t} \ \dots \ \rho_{nt} \ \beta_t \ \gamma_t]'$. Then:

$$\Phi_t = \Phi_{t-1} + \omega_t, \quad (8)$$

with

$$\omega_t \sim N(0, H^{-1}),$$

where $E(\varepsilon_t^* \omega_{is}) = 0$, for all $t, s = 1, \dots, T$, and $i = 1, \dots, n+3$. Furthermore the driftless random walks for the coefficients of the lags of the dependent variable are subject to *reflecting barriers* that ensure that the roots of the polynomial $1 - \rho_{1t}L - \rho_{2t}L^2 - \dots - \rho_{nt}L^n$ lie outside the unit circle⁴⁵.

The estimation of this model is achieved by augmenting the Gibbs sampling algorithm for the

⁴⁵The reflecting barriers are implemented in the MCMC algorithm with rejection sampling. The details are provided in the Appendix.

benchmark model with additional blocks for the Time-Varying Parameters and for the variance-covariance matrix of ω_t . Specifically, the additional blocks are:

$$\lambda_6 = \{H\}$$

$$\lambda_{7_{1,\dots,T}} = \{\Phi_t\}, t = 1, \dots, T$$

Note that, in contrast to the smoothing algorithm needed for the generation of the latent variables, the smoothing algorithm needed for generating the Φ_t 's is standard, as the model for the Time Varying Parameters can be cast into a state-space form with the Measurement Equation being the latent equation (equation (2''')), and with the Transition Equation being the driftless random walk equation for Φ_t (equation (8)), and then the Kalman filter can be used to obtain the conditional posteriors for $\Phi_t, t = 1, \dots, T$. The Appendix provides the details of how this smoothing algorithm and the rest of the Gibbs sampler for the TVP are implemented.

Four TVP models are estimated: One for the pre-Volcker period, one for the Volcker-Greenspan period, one for the whole sample, and an additional one for the whole sample that allows for the possibility of a jump⁴⁶ in the Time Varying Parameters in August of 1979 (Paul Volcker's appointment): The motivation for the last model is that if there is a discrete jump at that date, then a TVP model that does not allow for such a jump will be misspecified, and will thus potentially confound important aspects of the dynamics.

The results from the estimation of these 4 models are presented in Figures 4A-D⁴⁷. There is clear evidence from all these models of a sharp contrast in the inflation coefficient between the 1970's and the 1980's: The inflation coefficient is substantially lower (and insignificant too) in the pre-Volcker period than during the 1980's, and the full-sample results provide evidence of a relatively rapid increase in the inflation coefficient that begins in the late 1970's and culminates in the 1980's.

⁴⁶The jump is implemented by drawing the Time Varying Parameters of that date from a conditional posterior with a higher variance. That higher variance is exogenously specified, and I tried different values, with the results remaining quite similar in all these checks.

⁴⁷These figures also contain 90% bands around the estimated parameters. These bands are based on the quantiles of the conditional posterior distributions of the Time Varying Parameters.

These results are of course consistent with the policy mistakes hypothesis of CGG and provide further support to the results obtained with the benchmark model in split sample. There is also some evidence of an output gap coefficient that is higher and more significant more often during the 1970's than during the 1980's.

An interesting aspect of these results that is concealed by the split sample approach is the drop in the inflation coefficient that occurs during the Greenspan period: The 1990's were a rosy era of low inflation and low inflationary expectations, overall stability and high growth and these results may be revealing aspects of a monetary policy that reflected these realities of the 1990's.

6 Concluding remarks

This paper provides evidence that support the "policy mistakes" hypothesis as an explanation of the Great Inflation and of the stark contrast in US macroeconomic performance between the 1970's and the 1980's-1990's: There is strong evidence that prior to Paul Volcker's appointment as Chairman of the Federal Reserve in August of 1979, the Fed was too weak in its response to anticipated inflation and that under Volcker the Fed adopted a much stronger anti-inflationary stance. The results also suggest that the pre-Volcker Fed was too activist in its response to real-time estimates of the output gap that were upwardly biased, while starting with Volcker the Fed adopted a more cautious approach towards such real-time estimates.

While the Fed's excessive responsiveness to inaccurate real-time estimates of the output gap during the 1970's probably contributed to the instability of that period, it is clear that, as Clarida, Galí, and Gertler (2000) claim, the Fed's stance towards inflation, which was accommodative prior to Volcker's appointment, and which shifted dramatically towards strong anti-inflationary policies under Volcker, was a central factor behind the Great Inflation episode of the 1970's and the period of low inflation and macroeconomic stability that followed.

These results are not necessarily incompatible with alternative theories, such as ones that emphasize the role of non-policy shocks, or Time Varying Parameter approaches that suggest that there were richer dynamics or that changes in policy were more gradual than suggested by CGG.

Indeed, the central results and conclusions of this paper are preserved when a series of robustness checks and extensions that incorporate features of these alternative theories are considered.

Thus, these results provide a motivation and empirical support for learning models where the monetary policy-making authority initially has incorrect perceptions in real time about the structure of the economy and eventually corrects such misperceptions through accumulated knowledge coming from past policy mistakes and from the experience of the impact of these mistakes on the economy.

Appendix: MCMC Algorithms for the Benchmark Model and for its Extensions

A few words on notation and terminology for what follows first: θ will be taken to mean in what follows all the variables *other than the ones being generated in the particular block under consideration*. Furthermore, ff , ff^* , X will denote the entire vector for the dependent variable, the latent dependent variable, and the entire matrix of explanatory variables, respectively, (periods $1, \dots, T$), and ff_t , ff_t^* , X_t will denote the dependent variable, the latent dependent variable, and the explanatory variables for period t , respectively. The word "conditional" will be taken to mean *conditional on everything, except of course for the variable(s) being generated in the particular block under consideration*.

Algorithm for the Benchmark Model:

Generating the Variance: Inverted gamma distributions are convenient priors for the variance, since when multiplied by the conditional likelihood, they result in conditional posteriors which are also inverted gammas⁴⁸, that we know how to sample from:

So, if the prior for σ^2 is $IG(\frac{\nu_0}{2}, \frac{\delta_0}{2})$, where IG stands for inverted gamma, then the conditional posterior is also $IG(\frac{\nu_1}{2}, \frac{\delta_1}{2})$, where $\nu_1 = \nu_0 + T$, and $\delta_1 = \delta_0 + \varepsilon^{*'}\varepsilon^*$, where ε^* is the $T \times 1$ vector of latent error terms of equation (2).

I conducted experiments using both flat priors⁴⁹ and various forms of inverted gamma priors (I let each of the two parameters of the Gamma pdf vary from 0.1 to 10 and I tried various combinations of the two parameters within that range). The results were quite similar in all these experiments.

Generating the coefficients of the explanatory variables: A flat prior for this block results in a Gaussian conditional posterior from which I can sample easily: In particular, this conditional

⁴⁸See, for instance, Kim & Nelson (1999) for the derivation of this.

⁴⁹An appropriate flat prior for the variance is the positive half of the real line and it results in a posterior for the variance which is proportional to the conditional likelihood for the variance (thus it's proportional to $(\sigma^2)^{-\frac{T}{2}} \times e^{-\frac{1}{2\sigma^2}\varepsilon^{*'}\varepsilon^*}$) and whose support is the positive real line. This is also of the Gamma form.

posterior is, in a standard way, $N((X'X)^{-1}X'ff^*, \sigma^2(X'X)^{-1})$. For a derivation of this, see, for instance, Albert and Chib (1993b). I also conducted robustness checks using proper conjugate Gaussian priors, and the results did not change significantly.

The required stationarity constraints on the coefficients of the lags of the latent dependent variable are implemented with rejection sampling, whereby draws from the posterior for the coefficients are taken until the constraints are satisfied, (and the draws are discarded when they do not satisfy the constraints).

Generating the latent dependent variables: I use a single-move smoothing algorithm here, which entails simulating each ff_t^* , $t = 1, \dots, T$, one by one in separate blocks, while also conditioning on all the data, and all the other parameters, including all the other latent variables, for each block. The algorithm is derived as follows:

Let $g(ff_t^*|\theta, ff, X)$ denote the conditional distribution of ff_t^* , and let \widetilde{ff}_t^* denote all the latent variables for periods $1, \dots, t$, and let $\widetilde{ff}_{\neq t}^*$ denote all the latent variables *for all periods except for t* , and similarly let \widetilde{ff}_t denote all the dependent variables for periods $1, \dots, t$. The dependence on the parameters other than the latent variables and on the explanatory variables is suppressed in what follows for convenience. Furthermore, for expositional purposes, I present the case of one lag for the latent variable. The proof for more than one lags is the same. So, we have that:

$$\begin{aligned}
g(ff_t^*|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_T^*) &= g(ff_t^*|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t, ff_{t+1}, \dots, ff_T) \\
&= \frac{g(ff_t^*, ff_{t+1}, \dots, ff_T|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t)}{g(ff_{t+1}, \dots, ff_T|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t)} \\
&= \frac{g(ff_t^*|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t)g(ff_{t+1}, \dots, ff_T|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t, ff_t^*)}{g(ff_{t+1}, \dots, ff_T|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t)} \\
&= g(ff_t^*|\widetilde{ff}_{\neq t}^*, \widetilde{ff}_t) \\
&= g(ff_t^*|\widetilde{ff}_{t-1}^*, ff_{t+1}^*, \dots, ff_T^*, \widetilde{ff}_{t-1}, ff_t) \\
&= \frac{g(ff_t^*, ff_{t+2}^*, \dots, ff_T^*|\widetilde{ff}_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_{t-1}, ff_t)}{g(ff_{t+2}^*, \dots, ff_T^*|\widetilde{ff}_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_{t-1}, ff_t)} \\
&\propto g(ff_t^*, ff_{t+2}^*, \dots, ff_T^*|\widetilde{ff}_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_{t-1}, ff_t) \\
&= g(ff_t^*|\widetilde{ff}_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_{t-1}, ff_t)g(ff_{t+2}^*, \dots, ff_T^*|\widetilde{ff}_{t-1}^*, ff_t^*, ff_{t+1}^*, \widetilde{ff}_{t-1}, ff_t) \\
&\propto g(ff_t^*|\widetilde{ff}_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_{t-1}, ff_t) \\
&= g(ff_t^*|ff_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_t).
\end{aligned}$$

Note that the transition from the 3rd line to the 4th line is valid as ff_{t+1}, \dots, ff_T do not depend on ff_t^* , given $\widetilde{ff}_{\neq t}^*$. Note also that the transition from the 5th line to the 6th line is valid (for the case of models with one lag for the latent variable) because the denominator of the fraction of the 5th line does not depend on ff_t^* .

The pdf of the resulting distribution, namely $g(ff_t^*|ff_{t-1}^*, ff_{t+1}^*, \widetilde{ff}_t)$ can be obtained from the joint distribution of all the error terms where ff_t^* appears. For the case with one latent lag, ff_t^* appears in the equations giving ε_t^* , and ε_{t+1}^* . The joint pdf of the error terms is Gaussian, and ignoring for a moment the effect of conditioning on \widetilde{ff}_t , it is easy to show⁵⁰ that ff_t^* is distributed (given ff_{t-1}^*, ff_{t+1}^*) as $N(\frac{\beta'X_t + \beta_1 ff_{t-1}^* + \beta_1(ff_{t+1}^* - \beta'X_{t+1})}{1 + \beta_1^2}, \sigma^2)$, where β, X_t, X_{t+1} are defined here to exclude the latent lag and its coefficient, and β_1 is the coefficient of the latent lag.

The effect of conditioning on \widetilde{ff}_t is a truncation, and the form of the truncation is determined

⁵⁰Just rewrite, in that joint pdf, each of the error terms as an expression of the latent variables that appear in the latent equation that corresponds to that error term.

by equation (3):

If $\Delta f f_t \in \text{category } j$, then $f f_t^* \in (f f_{t-1} + c_{j-1}, f f_{t-1} + c_j), \forall j, \forall t$.

Thus, the required sampling task is that of sampling from a (univariate) truncated normal. The best way of doing that is a combination of sampling from a uniform and inverting the truncated normal cdf. Specifically, I wish to simulate the latent dependent variable, which has a Normal cdf F with mean μ and variance σ^2 , but that is truncated between a and b . Let $Z \sim \text{uniform}(0, 1)$. Then $F^{-1}(Z) \sim F$. Therefore, and since $F(f f_t^*) = \frac{\Phi(\frac{f f_t^* - \mu}{\sigma}) - \Phi(\frac{a - \mu}{\sigma})}{\Phi(\frac{b - \mu}{\sigma}) - \Phi(\frac{a - \mu}{\sigma})}$, where Φ is the standard normal cdf, I simulate the latent dependent variable by sampling from: $\sigma \Phi^{-1}\{Z[\Phi(\frac{b - \mu}{\sigma}) - \Phi(\frac{a - \mu}{\sigma})] + \Phi(\frac{a - \mu}{\sigma})\} + \mu$.

Additional blocks needed for the threshold coefficients:

Here I can adopt a diffuse prior for the threshold coefficients as then the conditional posterior becomes a uniform distribution (that we know how to sample from). In particular, it can be shown (see Albert and Chib (1993b)) that the conditional posterior for $c_j, j = 1, \dots, J$, (where the conditioning is on the other threshold coefficients too, in addition to everything else) is:

$$\text{uniform}[\max\{\max\{f f_t^* : \Delta f f_t = j\}, c_{j-1}\}, \min\{\min\{f f_t^* : \Delta f f_t = j + 1\}, c_{j+1}\}].$$

Additional blocks needed to implement the regime-switching extension:

Generating the State Variables: The *states*, S_1, \dots, S_T are Bernoulli random variables. Let $\widetilde{S}_T = [S_1 \dots S_T]$. Let Y_1, \dots, Y_T denote the data for periods $1, \dots, T$, respectively, and let \widetilde{Y}_t denote all the data up to period $t, t = 1, \dots, T$. Let $g(\widetilde{S}_T | \theta, \widetilde{Y}_T)$ denote the conditional distribution of \widetilde{S}_T . Following Kim and Nelson (1999), I adopt the *Multi-Move Gibbs sampling* approach; that is, I draw all of the states together in a single block, that is, from $g(\widetilde{S}_T | \theta, \widetilde{Y}_T)$, which, as Kim and Nelson (1999) demonstrate can be simplified as follows:

$$\begin{aligned} g(\widetilde{S}_T | \theta, \widetilde{Y}_T) &= g(S_1, \dots, S_T | \theta, \widetilde{Y}_T) \\ &= g(S_T | \theta, \widetilde{Y}_T) g(S_{T-1}, \dots, S_1 | S_T, \theta, \widetilde{Y}_T) \\ &= g(S_T | \theta, \widetilde{Y}_T) g(S_{T-1} | S_T, \theta, \widetilde{Y}_T) g(S_{T-2}, \dots, S_1 | S_T, S_{T-1}, \theta, \widetilde{Y}_T) \\ &= g(S_T | \theta, \widetilde{Y}_T) g(S_{T-1} | S_T, \theta, \widetilde{Y}_T) g(S_{T-2} | S_T, S_{T-1}, \theta, \widetilde{Y}_T) \dots g(S_1 | S_{T-1}, \dots, S_2, \theta, \widetilde{Y}_T) \end{aligned}$$

$$\begin{aligned}
&= g(S_T|\theta, \widetilde{Y}_T)g(S_{T-1}|S_T, \theta, \widetilde{Y}_{T-1})g(S_{T-2}|S_{T-1}, \theta, \widetilde{Y}_{T-2})\dots g(S_1|S_2, Y_1) \\
&= g(S_T|\theta, \widetilde{Y}_T) \prod_{t=1}^{T-1} g(S_t|S_{t+1}, \theta, \widetilde{Y}_t).
\end{aligned}$$

The simplifications above occur because of the Markov property of the states. So, and as suggested by the last expression, we can generate first S_T conditional on θ, \widetilde{Y}_T , and then for $t = T-1, \dots, 1$, we generate S_t conditional on \widetilde{Y}_t and S_{t+1} . So, I run Hamilton's filter (Hamilton (1989)) to obtain, for all t , $g(S_t|\theta, \widetilde{Y}_t)$, and $g(S_T|\theta, \widetilde{Y}_T)$ in particular, from which I generate S_T . Then I generate the states of the previous periods using the following:

$$g(S_t|\theta, \widetilde{Y}_t, S_{t+1}) = \frac{g(S_t, S_{t+1}|\theta, \widetilde{Y}_t)}{g(S_{t+1}|\theta, \widetilde{Y}_t)} = \frac{g(S_{t+1}|S_t, \theta, \widetilde{Y}_t)g(S_t|\theta, \widetilde{Y}_t)}{g(S_{t+1}|\theta, \widetilde{Y}_t)}.$$

Then, using this expression, I calculate $\Pr[S_t = 1|\theta, S_{t+1}, \widetilde{Y}_t]$ as follows:

$$\Pr[S_t = 1|\theta, S_{t+1}, \widetilde{Y}_t] = \frac{g(S_{t+1}|S_t=1, \theta)g(S_t=1|\theta, \widetilde{Y}_t)}{\sum_{j=0}^1 g(S_{t+1}|S_t=j, \theta)g(S_t=j|\theta, \widetilde{Y}_t)}.$$

To determine whether S_t is then 0 or 1, I take a draw from the uniform distribution between 0 and 1, and if the generated number is less than or equal to this probability I set $S_t = 1$. Otherwise I set $S_t = 0$.

Generating the Transition Probabilities: The beta distribution (used here) is a convenient conjugate prior for the transition probabilities, as it is easy to sample from a beta distribution. The priors for q and p are: $q \sim \text{beta}(u_{00}, u_{01})$, $p \sim \text{beta}(u_{11}, u_{10})$,

where I set the hyperparameters of the priors equal to different values and I observe that the results are quite similar across these experiments.

The conditional likelihood is: $L(q, p) = q^{n_{00}}(1 - q)^{n_{01}}p^{n_{11}}(1 - p)^{n_{10}}$,

where n_{ij} is equal to the number of transitions from state i to state j , $i, j = 0, 1$.

The resulting posteriors are also beta distributions⁵²:

$$q \sim \text{beta}(u_{00} + n_{00}, u_{01} + n_{01}), p \sim \text{beta}(u_{11} + n_{11}, u_{10} + n_{10}).$$

Generating the Variances: As with the benchmark model, the priors used here are inverted gammas:

⁵¹ $g(S_{t+1}|S_t)$ is the transition probability (generated in separate blocks), and $g(S_t|\theta, \widetilde{Y}_t)$, $g(S_{t+1}|\theta, \widetilde{Y}_t)$ have been obtained from Hamilton's filter.

⁵²See Kim and Nelson (1999) for a derivation of this result.

To generate σ_0^2 conditional on h_1 , I divide both sides of the latent equation by $\sqrt{1 + h_1 S_t}$, and it can be shown that if the conditional prior for σ_0^2 is $IG(\frac{\nu_0}{2}, \frac{\delta_0}{2})$, then the conditional posterior is also $IG(\frac{\nu_1}{2}, \frac{\delta_1}{2})$, where $\nu_1 = \nu_0 + T$, and $\delta_1 = \delta_0 + \sum_{t=1}^T (residuals^*)^2$, where $residuals^*$ are equal to the normalized LHS of the latent equation minus the normalized RHS of the latent equation.

To generate h_1 (and thus σ_1^2) conditional on σ_0^2 , I divide both sides of the latent equation by σ_0 , and, as before, it can be shown that if the conditional prior for h_1 is $IG(\frac{\nu_3}{2}, \frac{\delta_3}{2})$, then the conditional posterior is also $IG(\frac{\nu_4}{2}, \frac{\delta_4}{2})$, where $\nu_4 = \nu_3 + T^*$, and $\delta_4 = \delta_3 + \sum_{t=1}^{T^*} (residuals^*)^2$, where $residuals^*$ are as before (where now the division is done by σ_0), and where T^* is equal to the number of periods during which we are at the high volatility state.

As with the single variance case, I conducted experiments using both flat priors and various forms of inverted gamma priors (I let each of the two parameters of the Gamma pdf vary from 0.1 to 10 and I tried various combinations of the two parameters within that range). The results were quite similar in all these experiments.

The normalization constraint that $h_1 > 0$, which means that σ_1^2 is constrained to be the high state variance and thus that it must be higher than σ_0^2 (that is constrained to be the low state variance), is implemented with rejection sampling.

Additional blocks needed for the Time Varying Parameters:

Generating the Time Varying Parameters: In contrast to the case of the latent variables, obtaining the smoothing algorithm for the Time Varying Parameters is standard because here we can usefully employ a state-space representation, with the Measurement Equation being the latent equation (equation (2''')), and with the Transition Equation being the driftless random walk for the TVP's (equation (8)), together with the Kalman filter.

Specifically, let $\tilde{\Phi}_T = [\Phi_1 \dots \Phi_T]'$. Let Y_1, \dots, Y_T denote the data for periods 1, ..., T , respectively, and let \tilde{Y}_t denote all the data up to period $t, t = 1, \dots, T$. Let $g(\tilde{\Phi}_T | \theta, \tilde{Y}_T)$ denote the conditional distribution of $\tilde{\Phi}_T$. Then, following Kim and Nelson (1999) I employ a multimove Gibbs-sampling approach, thus generating the entire $\tilde{\Phi}_T$ as a block from its conditional distribution, $g(\tilde{\Phi}_T | \theta, \tilde{Y}_T)$.

The Markov property of the Φ_t 's ensure that convenient simplifications occur in $g(\widetilde{\Phi}_T|\theta, \widetilde{Y}_T)$, and in particular:

$$\begin{aligned}
g(\widetilde{\Phi}_T|\theta, \widetilde{Y}_T) &= g(\Phi_T|\theta, \widetilde{Y}_T)g(\widetilde{\Phi}_{T-1}|\theta, \Phi_T, \widetilde{Y}_T) \\
&= g(\Phi_T|\theta, \widetilde{Y}_T)g(\Phi_{T-1}|\theta, \Phi_T, \widetilde{Y}_T)g(\widetilde{\Phi}_{T-2}|\theta, \Phi_{T-1}, \Phi_T, \widetilde{Y}_T) \\
&= \dots \\
&= g(\Phi_T|\theta, \widetilde{Y}_T)g(\Phi_{T-1}|\theta, \Phi_T, \widetilde{Y}_T)g(\Phi_{T-2}|\theta, \Phi_{T-1}, \widetilde{Y}_T)\dots g(\Phi_1|\theta, \Phi_2, \widetilde{Y}_T) \\
&= g(\Phi_T|\theta, \widetilde{Y}_T)g(\Phi_{T-1}|\theta, \Phi_T, \widetilde{Y}_{T-1})g(\widetilde{\Phi}_{T-2}|\theta, \Phi_{T-1}, \widetilde{Y}_{T-2})\dots g(\Phi_1|\theta, \Phi_2, \widetilde{Y}_1) \\
&= g(\Phi_T|\theta, \widetilde{Y}_T) \prod_{t=1}^{T-1} g(\Phi_t|\theta, \Phi_{t+1}, \widetilde{Y}_t)
\end{aligned}$$

As suggested by this last expression, I first need to generate Φ_T from $g(\Phi_T|\theta, \widetilde{Y}_T)$, and then, given Φ_{t+1} , generate Φ_t from $g(\Phi_t|\theta, \Phi_{t+1}, \widetilde{Y}_t), t = \dots, T - 1$. Thus, I first generate Φ_T from $g(\Phi_T|\theta, \widetilde{Y}_T) \sim N(\Phi_{T|T}, P_{T|T})$, and then Φ_t , for $t = T-1, \dots, 1$ from $g(\Phi_t|\theta, \Phi_{t+1}, \widetilde{Y}_t) \sim N(\Phi_{t|t, \Phi_{t+1}}, P_{t|t, \Phi_{t+1}})$, where $\Phi_{T|T} = E(\Phi_T|\theta, \widetilde{Y}_T)$, $P_{T|T} = Cov(\Phi_T|\theta, \widetilde{Y}_T)$, $\Phi_{t|t, \Phi_{t+1}} = E(\Phi_t|\theta, \widetilde{Y}_t, \Phi_{t+1}) = E(\Phi_t|\theta, \Phi_{t+1})$, $P_{t|t, \Phi_{t+1}} = Cov(\Phi_t|\theta, \widetilde{Y}_t, \Phi_{t+1}) = Cov(\Phi_t|\theta, \Phi_{t+1})$. The updating terms $\Phi_{T|T}, P_{T|T}$, (and also all $\Phi_{t|t}, P_{t|t}, t = 1, \dots, T$) can be derived in a standard way using the Kalman filter⁵³. The same holds true for the terms $\Phi_{t|t, \Phi_{t+1}}$, and $P_{t|t, \Phi_{t+1}}$ since they can also be viewed as updating terms in which the updating is done not with Y_t , but with Φ_{t+1} , which has been generated, and thus can be considered as observed data.

The initial values, $\beta_{0|0}$ are arbitrary, with $P_{0|0}$ having large diagonal elements (so that large uncertainty is attached to $\beta_{0|0}$).

The reflecting barriers imposing the stability condition on the coefficients of the lags of the dependent variable are implemented with rejection sampling, done separately for each time period $t = 1, \dots, T$.

Generating the precision matrix H : The prior for H is Wishart, $W(\nu_0, H_0)$, where I set $\nu_0 = 0, H_0^{-1} = 0$, and then the conditional posterior for H is also Wishart, $W(\nu_1, H_1)$, where $\nu_1 = T + \nu_0, H_1 = [H_0^{-1} + \sum_{t=1}^T (\Phi_t - \Phi_{t-1})(\Phi_t - \Phi_{t-1})']^{-1}$.

⁵³See Hamilton (1994), and Kim and Nelson (1999).

A note on the computer code: All of the computer code for this paper was written in Gauss, Version 3.2.34. The seed was always fixed at 180303.

References

- Albert, J., and S. Chib (1993a): "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business and Economic Statistics*, January, Vol., 11, No. 1, pp. 1-15.
- Albert, J., and S. Chib (1993b): "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, June, Vol., 88, No. 422, pp. 669-679.
- Bernanke, B. S., and A. S. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82, 901-21.
- Bernanke, B., and I. Mihov (1998): "Measuring Monetary Policy," *Quarterly Journal of Economics* 113(3), 869-902.
- Bernstein, S. (1917): *Theory of Probability* (in Russian), 4th Edition (1946) Gostekhizdat, Moscow-Leningrad.
- Blanchard, O.J., and J. Simon (2001): "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity*, 2001-1, 135-174.
- Boivin, J. (2004): "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," *mimeo, Columbia*.
- Chernozhukov, V., and H. Hong (2003): "An MCMC approach to classical estimation," *Journal of Econometrics* (115), 293-346.
- Chib, S. (2001): "Markov Chain Monte Carlo Methods: Computation and Inference." In: Heckman, J., and Leamer, E. (Eds.), *Handbook of Econometrics*, Vol. 5, North-Holland, Amsterdam, p. 3569-3649.
- Clarida, R., Galí, J., and M. Gertler (1998): "Monetary Policy Rules in Practice: Some International Evidence," *European Economic Review* 42: 1033-1067.

Clarida, R., Galí, J., and M. Gertler (1999): "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature* Vol XXXVII, No. 4: 1661-1707.

Clarida, R., Galí, J., and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics* 115(1), 147-180.

Cogley, T., and T. Sargent (2001): "Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual* 16, pp. 331-373.

Cogley, T., and T. Sargent (2002): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *mimeo, New York University*.

Cook, T., and T. Hahn (1989): "The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970's," *Journal of Monetary Economics* 24: 331-351.

Dueker, Michael (1999a): "Measuring Monetary Policy Inertia in Target Fed Funds Rate Changes," *Federal Reserve Bank of St. Louis Review* 81(5): 3-9.

Dueker, Michael (1999b): "Conditional Heteroskedasticity in Qualitative Response Models of Time Series: A Gibbs Sampling Approach to the Bank Prime Rate," *Journal of Business and Economic Statistics* 17(4): 466-472.

Eichengreen, B., M. Watson, and R. Grossman (1985): "Bank Rate Policy under the Interwar Gold Standard: A Dynamic Probit Model," *Economic Journal* 95: 725-745.

Geweke, John (1991): "Efficient Simulation from the Multivariate Normal and Student Distributions subject to Linear Constraints," *Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface*, 571-578.

Geweke, J., and M. Keane (2001): "Computationally Intensive Methods for Integration in Econometrics." In: Heckman, J, and Leamer, E. (Eds.), *Handbook of Econometrics, Vol. 5*, North-Holland, Amsterdam, p. 3463-3568.

- Goodfriend, M. (1991): "Interest Rates and the Conduct of Monetary Policy," *Carnegie-Rochester Conference Series on Public Policy*, XXXIV, 7-30.
- Hajivassiliou, Vassilis (1990): "Smooth Simulation Estimation of Panel Data LDV Models", *unpublished*.
- Hajivassiliou, V., and D. McFadden (1998): "The Method of Simulated Scores for the Estimation of LDV Models," *Econometrica* 66(4), 863-96.
- Hamilton, James D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357-384.
- Hamilton, James D. (1994): *Time Series Analysis*, Princeton University Press, Princeton.
- Hamilton, J., and O. Jordà (2002): "A Model for the Federal Funds Rate Target," forthcoming, *Journal of Political Economy*.
- Hamilton, J., Waggoner, D., and T. Zha (2003): "Normalization in Econometrics," *manuscript, UCSD*.
- Jalil, M. (2004): "Monetary Policy in retrospective: a Taylor rule inspired exercise," *mimeo, UCSD*.
- Judd, J., and G. Rudebusch (1998): "Taylor's Rule and the Fed: 1970-1997," *Federal Reserve Bank of San Francisco Economic Review* Number 3, 3-16.
- Judd, J., and G. Rudebusch (1998): "Taylor's Rule and the Fed: 1970-1997," *Federal Reserve Bank of San Francisco Economic Review* Number 3, 3-16.
- Keane, Michael (1994): "A computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 62(1), 95-116.
- Kim, C., and C. Nelson (1999a): *State-Space Models With Regime Switching*, MIT Press, Cambridge.

- Kim, C., and C. Nelson (1999b): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of Business Cycle," *The Review of Economics and Statistics*, 81, 608-616.
- Lee, Lung-fei (1999): "Estimation of Dynamic and ARCH Tobit Models", *Journal of Econometrics* 92: 355-390.
- Lehmann E. L., and G. Casella (1998): *Theory of Point Estimation*, 2nd edition, Springer.
- Lerman, S., and C. Manski (1981): "On the Use of Simulated Frequencies to Approximate Choice Probabilities." In: Manski, C., and McFadden, D. (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, p. 305-319.
- Mavroeidis, S. (2002): "Identification and Misspecification Issues in Forward-Looking Models," *UVA Econometrics Discussion Paper 2002/21*.
- Mavroeidis, S. (2004): "Weak Identification of Forward-looking Models in Monetary Economics," *Oxford Bulletin of Economics and Statistics*, 66, supplement: 609-635.
- McCulloch, R., and P. Rossi (1994): "An Exact Likelihood Analysis of the Multinomial Probit Model," *Journal of Econometrics* 64: 207-240.
- McFadden, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Choice Models without Numerical Integration," *Econometrica* 57, 995-1026.
- Meulendyke, Ann Marie (1998): *U.S. Monetary Policy and Financial Markets*. New York: Federal Reserve Bank of New York.
- Monokroussos, G. (2004): "A Classical MCMC Approach to the Estimation of Limited Dependent Variable Models of Time Series," *mimeo, UCSD*.
- Orphanides, A. (2001): "Monetary Policy Rules Based on Real-Time Data," *American Economic Review* 91(4), 964-985.

- Orphanides, A. (2002): "Monetary Policy Rules and the Great Inflation," *American Economic Review, Papers and Proceedings*, 92(2), 115-120, May 2002.
- Orphanides, A. (2004): "Monetary Policy Rules, Macroeconomic Stability and Inflation: A View from the Trenches," *Journal of Money, Credit and Banking*, 35(6).
- Pakes, A., and D. Pollard (1989): "Simulation and the asymptotics of optimization estimators," *Econometrica*, 57, 1027-1057.
- Piazzesi, Monika (2001): "An Econometric Model of the Yield Curve with Macroeconomic Jump Effects," *NBER Working Paper* No. 8246.
- Primiceri, G. (2004): "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy," *mimeo, Princeton*.
- Robert, C.P., and G. Casella (1999): *Monte Carlo Statistical Methods*, Springer.
- Rotemberg, J., and M. Woodford (1997): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," *NBER Macroeconomics Annual*, Bernanke, B., and J. Rotemberg, eds.
- Rotemberg, J., and M. Woodford (1999): "Interest Rate Rules in an Estimated Sticky Price Model," in *Monetary Policy Rules*, J. B. Taylor, ed.
- Rudebusch, G. (1995): "Federal Reserve Interest Rate Targeting, Rational Expectations and the Term Structure," *Journal of Monetary Economics* 35: 245-274.
- Sack, B., and W. Wieland (2000): "Interest Rate Smoothing and Optimal Monetary Policy: A Review of Recent Empirical Evidence," *Journal of Economics and Business* 52, 205-228.
- Sims, C. (1999): "Drift and Breaks in Monetary Policy," *mimeo, Princeton University*.
- Sims, C., and T. Zha (2004): "Were There Regime Switches in US Monetary Policy?," *mimeo, Princeton University*.

Stock, J. H., and M. W. Watson (2003): "Understanding Changes in International Business Cycle Dynamics," *mimeo*, Princeton University.

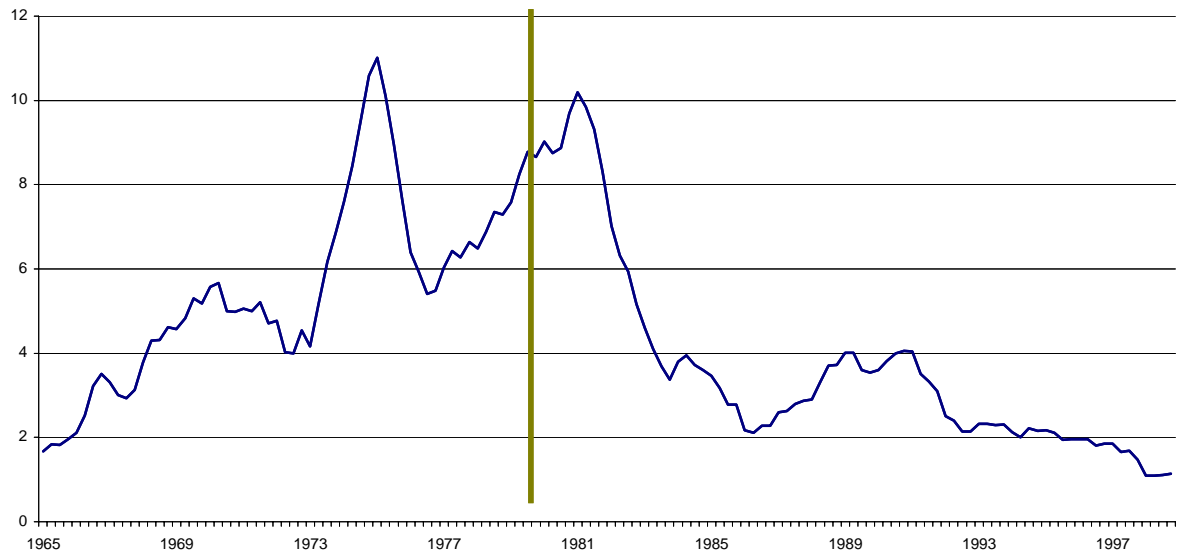
Taylor, John B. (1993): "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (December 1993), 195-214.

Taylor, J. B. (1999): *Monetary Policy Rules*, The University of Chicago Press.

Tanner, M.A., and W.H. Wong (1987): "The calculation of posterior distributions by data augmentation," *Journal of the American Statistical Association*, 82: 528-549.

von Mises, R. (1931), *Wahrscheinlichkeitsrechnung*, Springer-Verlag, Berlin.

Figure 1
U.S. Inflation



Note: Vertical line represents Paul Volcker's appointment as Chairman of the Federal Reserve in August of 1979.

Figure 2
Fed Funds Rate and its Target, FOMC Months

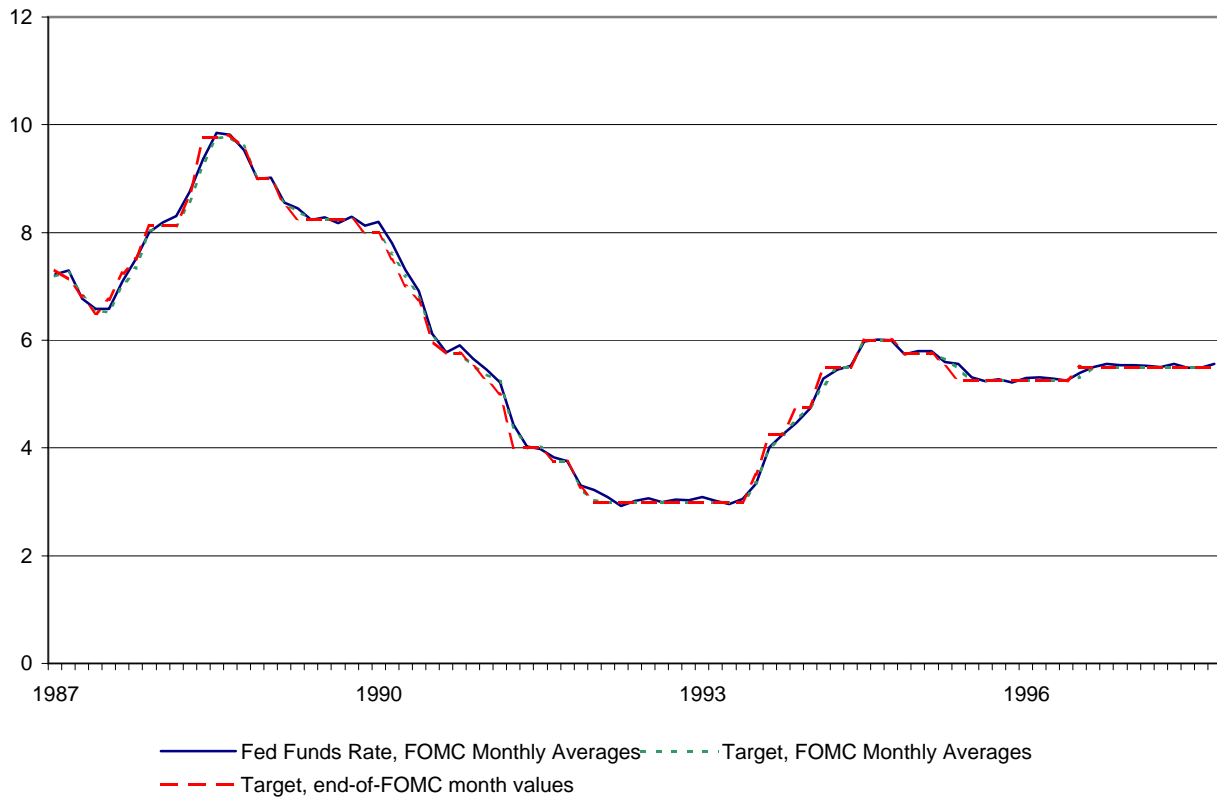


Table 1

Greenspan Period

<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>
<i>FOMC monthly averages of fed funds rate</i>									
<i>intercept</i>	0.0108	0.1555	0.0069	-0.2637	-0.2252	-0.1819	0.2037	0.2783	0.3406
<i>variance</i>	0.1593	0.0598	0.1466	0.0777	0.0836	0.0961	0.2408	0.2708	0.3061
<i>1st lag</i>	0.7544	0.0685	0.7627	0.5782	0.6292	0.6653	0.83	0.8475	0.8635
<i>inflation</i>	0.4808	0.1363	0.4648	0.2567	0.2838	0.3174	0.6544	0.7229	0.7952
<i>output gap</i>	0.4035	0.1006	0.3908	0.2401	0.2605	0.2879	0.5332	0.5947	0.647
<i>FOMC monthly averages of target</i>									
<i>intercept</i>	0.0256	0.159	0.0207	-0.2641	-0.2129	-0.172	0.2224	0.312	0.3579
<i>variance</i>	0.1659	0.0638	0.1526	0.0777	0.0863	0.0974	0.2512	0.2837	0.3098
<i>1st lag</i>	0.7542	0.0668	0.7657	0.5903	0.6228	0.6686	0.8287	0.844	0.8648
<i>inflation</i>	0.4754	0.1325	0.4617	0.2532	0.283	0.3204	0.6433	0.7049	0.78
<i>output gap</i>	0.4042	0.1003	0.3892	0.2422	0.2601	0.2907	0.5363	0.5938	0.6421
<i>Value of target at end of FOMC months</i>									
<i>intercept</i>	0.0182	0.1538	0.0124	-0.2609	-0.2161	-0.17	0.2112	0.271	0.3358
<i>variance</i>	0.1588	0.057	0.1471	0.0798	0.0885	0.0986	0.2356	0.2694	0.3039
<i>1st lag</i>	0.7658	0.0663	0.7733	0.6099	0.6465	0.676	0.842	0.8593	0.8732
<i>inflation</i>	0.4502	0.1335	0.4392	0.2253	0.2531	0.2931	0.6268	0.6737	0.7421
<i>output gap</i>	0.3983	0.1001	0.3879	0.2354	0.2567	0.2817	0.5221	0.5858	0.6358

Note: The last 7 columns are quantiles of the posterior distribution.

Table 2

<i>Pre-Volcker Period</i>									
<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>
<i>1:1000 iterations</i>									
<i>intercept</i>	0.9548	0.4496	0.9676	0.0755	0.2338	0.3988	1.5274	1.6921	1.8477
<i>variance</i>	1.1297	0.2374	1.0935	0.7517	0.7925	0.8518	1.4398	1.5615	1.6794
<i>1st lag</i>	0.8664	0.0489	0.8672	0.7623	0.7819	0.8027	0.9267	0.94	0.9641
<i>inflation</i>	0.0682	0.0808	0.0676	-0.0809	-0.0596	-0.033	0.1769	0.2038	0.2326
<i>output gap</i>	0.3006	0.1	0.3004	0.1111	0.1376	0.1729	0.4251	0.4666	0.4941
<i>501:1000 iterations</i>									
<i>intercept</i>	0.9451	0.4472	0.956	0.0531	0.2338	0.3838	1.4903	1.6664	1.8351
<i>variance</i>	1.1216	0.2364	1.0856	0.7403	0.7806	0.8314	1.4465	1.5747	1.6534
<i>1st lag</i>	0.8661	0.0474	0.8687	0.7623	0.7817	0.802	0.9261	0.9368	0.9526
<i>inflation</i>	0.0704	0.08	0.0691	-0.0716	-0.0484	-0.0272	0.1817	0.2045	0.2254
<i>output gap</i>	0.3014	0.0996	0.3009	0.1081	0.1394	0.1715	0.4317	0.4671	0.4906
<i>1:500 iterations</i>									
<i>intercept</i>	0.9645	0.4522	0.973	0.082	0.2288	0.4021	1.5374	1.7037	1.9593
<i>variance</i>	1.1378	0.2383	1.1039	0.7705	0.814	0.8656	1.437	1.5324	1.7086
<i>1st lag</i>	0.8668	0.0505	0.8666	0.7656	0.7838	0.8027	0.9282	0.9466	0.971
<i>inflation</i>	0.066	0.0815	0.0636	-0.0837	-0.0658	-0.0344	0.165	0.1936	0.2362
<i>output gap</i>	0.2997	0.1004	0.2999	0.1127	0.1337	0.1742	0.4242	0.4647	0.4951
<i>Volcker-Greenspan Period</i>									
<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>
<i>1:1000 iterations</i>									
<i>intercept</i>	-0.0439	0.279	-0.0281	-0.5669	-0.5025	-0.4088	0.3199	0.401	0.4955
<i>variance</i>	1.2317	0.196	1.2084	0.902	0.943	0.9919	1.4954	1.5877	1.6677
<i>1st lag</i>	0.7679	0.0435	0.7668	0.6826	0.6958	0.7117	0.823	0.8396	0.8535
<i>inflation</i>	0.4518	0.0865	0.4555	0.2683	0.3054	0.343	0.5552	0.5882	0.6256
<i>output gap</i>	0.1144	0.0874	0.1133	-0.0532	-0.0301	-0.0007	0.23	0.265	0.2875
<i>501:1000 iterations</i>									
<i>intercept</i>	-0.0378	0.2695	-0.0231	-0.5487	-0.4879	-0.3941	0.3163	0.4054	0.5055
<i>variance</i>	1.2171	0.1839	1.2005	0.8945	0.9412	0.9929	1.4652	1.5274	1.645
<i>1st lag</i>	0.7688	0.0424	0.7668	0.6882	0.6963	0.7165	0.8238	0.8408	0.8524
<i>inflation</i>	0.4487	0.0843	0.4532	0.27	0.3021	0.3393	0.5507	0.5825	0.6108
<i>output gap</i>	0.1115	0.0836	0.1119	-0.0532	-0.0353	0.0049	0.2224	0.2564	0.2908
<i>1:500 iterations</i>									
<i>intercept</i>	-0.05	0.2882	-0.0338	-0.5914	-0.5273	-0.4269	0.336	0.3983	0.4955
<i>variance</i>	1.2464	0.2066	1.225	0.9098	0.9516	0.9907	1.534	1.6198	1.6802
<i>1st lag</i>	0.767	0.0446	0.767	0.6745	0.6939	0.7103	0.8214	0.8369	0.8556
<i>inflation</i>	0.455	0.0886	0.4565	0.2683	0.3054	0.3482	0.561	0.5967	0.6458
<i>output gap</i>	0.1173	0.091	0.1138	-0.0523	-0.0296	-0.0034	0.2366	0.2695	0.2842

Note: Benchmark specification with 1 lag of the dependent variable.
 The last 7 columns are quantiles of the posterior distribution.

Figure 3A
Histograms of Posterior Distributions, Benchmark Model, Pre-Volcker Period

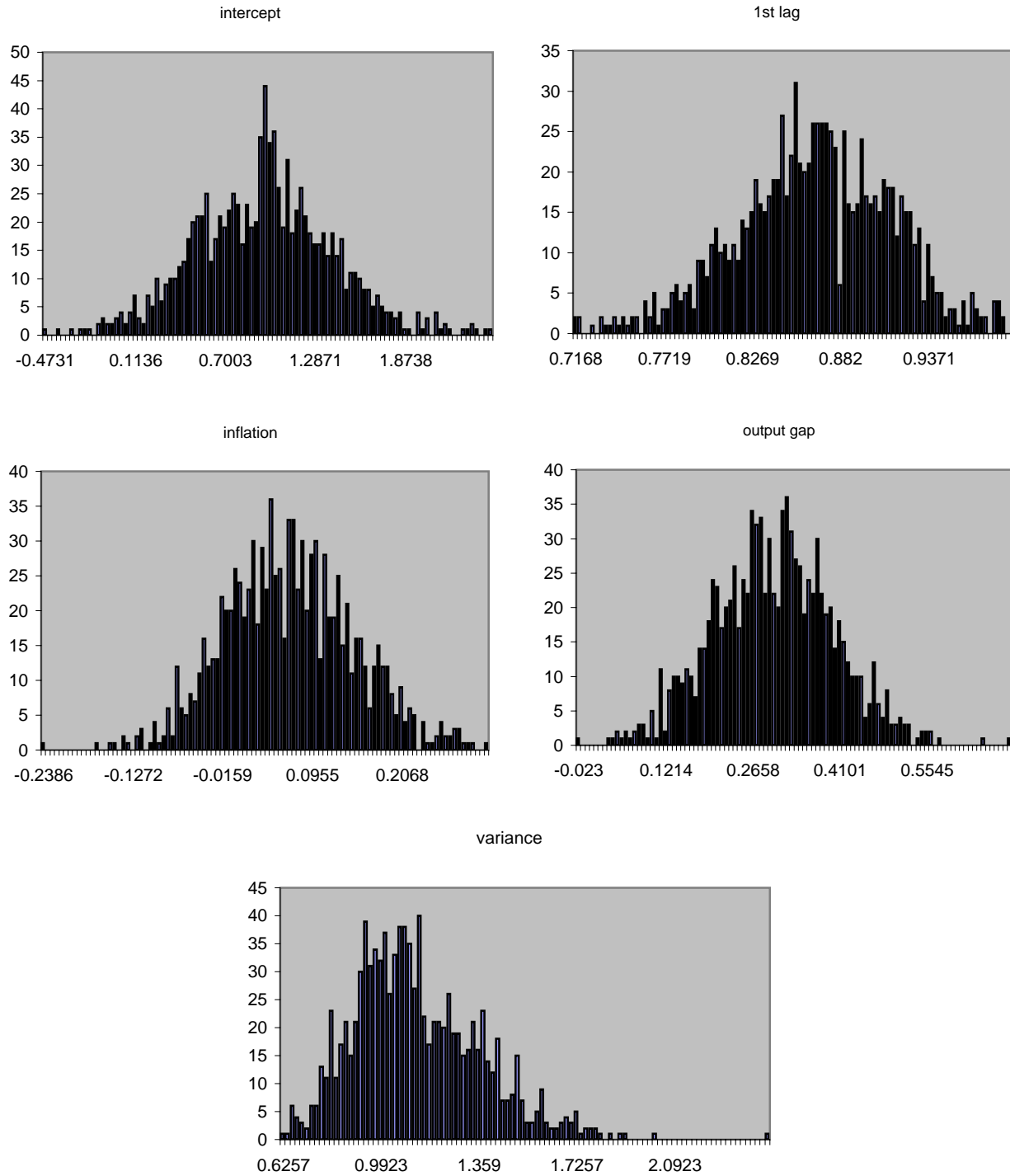


Figure 3B
Histograms of Posterior Distributions, Benchmark Model, Volcker-Greenspan Period

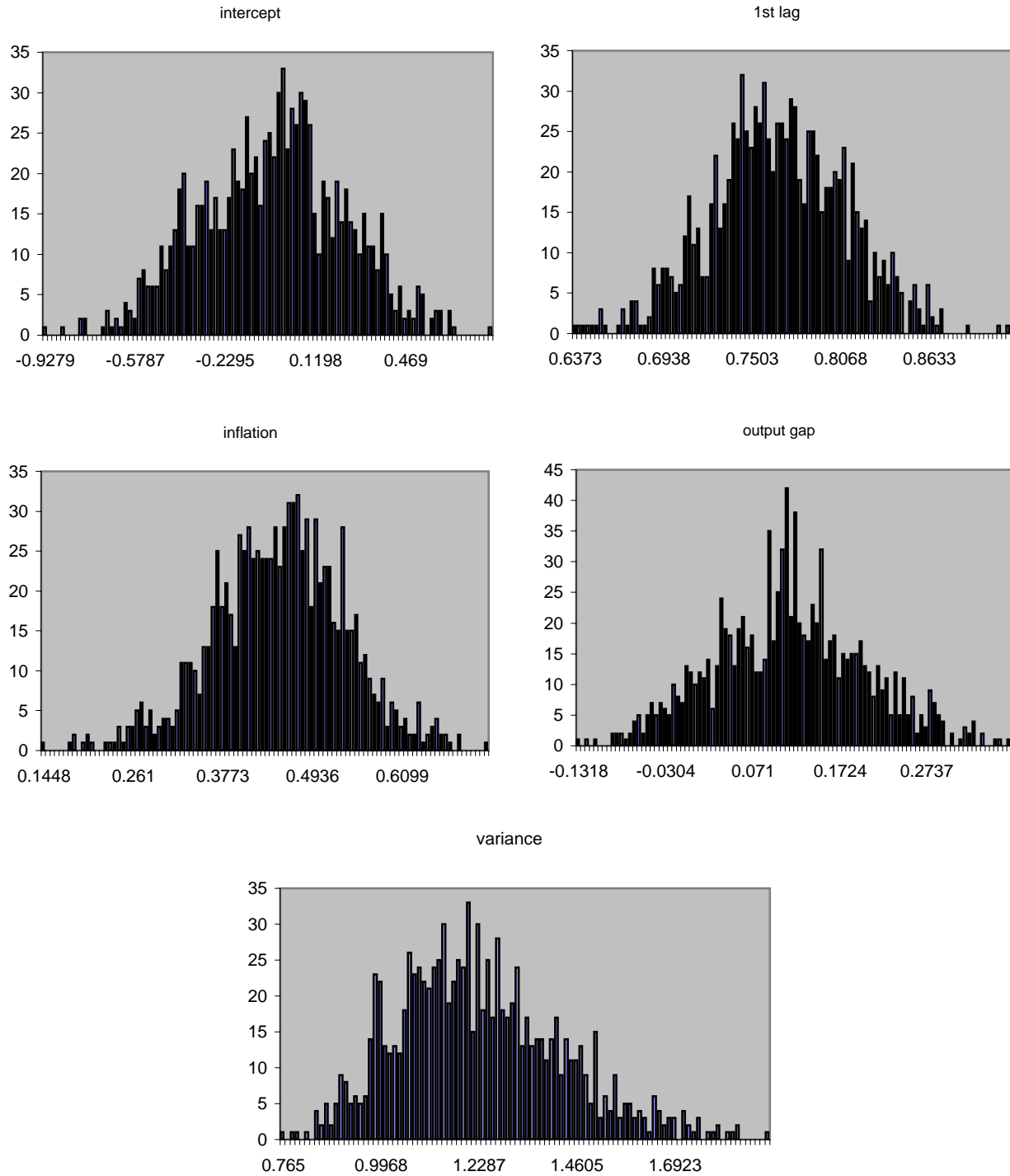


Table 3

Pre-Volcker Period									
variable	mean	std. dev.	median	2.5% qntl	5% qntl	10% qntl	90% qntl	95% qntl	97.5% qntl
1:1000 iterations									
intercept	0.9356	0.4392	0.9186	0.0721	0.2163	0.3811	1.4958	1.648	1.8073
variance	0.9956	0.2023	0.9776	0.6604	0.7078	0.7527	1.2621	1.3614	1.4329
1st lag	0.7471	0.1324	0.7437	0.5002	0.5298	0.5759	0.9192	0.9637	1.0185
2nd lag	0.1446	0.1315	0.1465	-0.1242	-0.0692	-0.0224	0.3136	0.3567	0.3852
inflation	0.0426	0.076	0.0407	-0.0988	-0.0825	-0.055	0.1366	0.1771	0.1975
output gap	0.3246	0.1004	0.3236	0.1236	0.1555	0.1917	0.4485	0.4899	0.5273
501:1000 iterations									
intercept	0.9581	0.4305	0.9564	0.1795	0.2535	0.4299	1.5292	1.6498	1.8073
variance	0.9812	0.1959	0.9586	0.6543	0.7013	0.7527	1.2472	1.3257	1.4036
1st lag	0.7505	0.1275	0.7524	0.498	0.5424	0.5749	0.9109	0.9594	1.0025
2nd lag	0.1428	0.1292	0.1446	-0.1025	-0.0692	-0.0203	0.3151	0.3574	0.3803
inflation	0.0368	0.0753	0.0358	-0.1053	-0.0866	-0.0623	0.1335	0.1755	0.1894
output gap	0.3283	0.0998	0.3295	0.1211	0.1555	0.1994	0.448	0.4899	0.5178
1:500 iterations									
intercept	0.9131	0.447	0.9114	0.0261	0.1619	0.348	1.4582	1.6307	1.8641
variance	1.0101	0.2078	1.0045	0.6685	0.7117	0.7527	1.2729	1.3871	1.4782
1st lag	0.7436	0.1372	0.7376	0.5015	0.5244	0.5759	0.9251	0.969	1.0376
2nd lag	0.1465	0.1338	0.1473	-0.1264	-0.0701	-0.0259	0.3131	0.3563	0.3979
inflation	0.0485	0.0764	0.0483	-0.0922	-0.0752	-0.0501	0.1414	0.1794	0.2163
output gap	0.3209	0.1011	0.3174	0.1316	0.1521	0.1857	0.452	0.4899	0.5291
Volcker-Greenspan Period									
variable	mean	std. dev.	median	2.5% qntl	5% qntl	10% qntl	90% qntl	95% qntl	97.5% qntl
1:1000 iterations									
intercept	-0.0455	0.2833	-0.0488	-0.5756	-0.4926	-0.41	0.3324	0.447	0.5006
variance	1.2293	0.2007	1.2099	0.8723	0.922	0.9869	1.4853	1.5926	1.6777
1st lag	0.7359	0.1059	0.7381	0.5237	0.5634	0.6067	0.8689	0.9145	0.9506
2nd lag	0.0299	0.0975	0.0318	-0.1635	-0.1332	-0.0927	0.1509	0.1866	0.2211
inflation	0.4595	0.0888	0.4606	0.276	0.3121	0.3479	0.5729	0.611	0.6369
output gap	0.1131	0.0872	0.1112	-0.0712	-0.037	-0.0022	0.2261	0.2586	0.2867
501:1000 iterations									
intercept	-0.0499	0.2899	-0.0472	-0.6045	-0.5237	-0.4235	0.3301	0.4455	0.4845
variance	1.2242	0.2008	1.2044	0.8587	0.9074	0.9731	1.4687	1.5845	1.6803
1st lag	0.731	0.1102	0.73	0.5068	0.5578	0.6003	0.8696	0.9218	0.9608
2nd lag	0.0359	0.1034	0.0453	-0.1707	-0.1484	-0.1101	0.1613	0.1958	0.2358
inflation	0.4596	0.0883	0.4623	0.2659	0.3075	0.3479	0.568	0.6061	0.6236
output gap	0.1157	0.0893	0.1137	-0.0722	-0.0343	0.0026	0.2343	0.2632	0.2878
1:500 iterations									
intercept	-0.0412	0.2767	-0.057	-0.5244	-0.4727	-0.398	0.3324	0.447	0.5662
variance	1.2344	0.2007	1.2181	0.8935	0.9291	1.0002	1.4963	1.5955	1.6705
1st lag	0.7409	0.1012	0.7468	0.5495	0.5791	0.6097	0.8688	0.9091	0.9421
2nd lag	0.0238	0.0909	0.023	-0.1462	-0.1221	-0.0856	0.1384	0.1713	0.2024
inflation	0.4595	0.0894	0.4559	0.2963	0.3183	0.347	0.5765	0.6129	0.6407
output gap	0.1105	0.085	0.11	-0.0706	-0.0387	-0.0067	0.2195	0.2484	0.276

Note: Benchmark specification with 2 lags of the dependent variable.
 The last 7 columns are quantiles of the posterior distribution.

Table 4

<i>Pre-Volcker Period</i>										
<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>	
<i>2-steps ahead</i>										
<i>intercept</i>	0.9719	0.4289	0.9627	0.1484	0.2933	0.4368	1.5111	1.6939	1.7968	
<i>variance</i>	1.0825	0.211	1.0622	0.724	0.7798	0.83	1.3447	1.4446	1.5276	
<i>1st lag</i>	0.806	0.0536	0.809	0.6962	0.71	0.7332	0.8717	0.89	0.9039	
<i>inflation</i>	0.1586	0.0745	0.1581	0.0105	0.0332	0.0628	0.2531	0.2817	0.3002	
<i>output gap</i>	0.3716	0.1047	0.3671	0.1773	0.2038	0.2365	0.5084	0.538	0.5776	
<i>1-step ahead</i>										
<i>intercept</i>	1.1139	0.4062	1.1045	0.3724	0.4543	0.5789	1.6488	1.8077	1.9074	
<i>variance</i>	0.9969	0.1949	0.9781	0.6548	0.7047	0.7624	1.2473	1.3441	1.408	
<i>1st lag</i>	0.7766	0.0585	0.7791	0.6607	0.6763	0.6998	0.8505	0.8732	0.8905	
<i>inflation</i>	0.168	0.0761	0.1683	0.0196	0.0435	0.0731	0.2649	0.2914	0.3118	
<i>output gap</i>	0.4182	0.1086	0.418	0.2109	0.242	0.2792	0.5578	0.5896	0.6281	
<i>0-steps ahead</i>										
<i>intercept</i>	1.5299	0.4589	1.5261	0.654	0.8093	0.9675	2.1268	2.3443	2.4859	
<i>variance</i>	1.0032	0.2014	0.9876	0.6515	0.7065	0.7603	1.2656	1.3821	1.4436	
<i>1st lag</i>	0.7423	0.0683	0.7433	0.6062	0.6241	0.6511	0.8274	0.8538	0.8688	
<i>inflation</i>	0.1288	0.0661	0.1284	0.0039	0.0228	0.045	0.2129	0.242	0.2676	
<i>output gap</i>	0.4764	0.1264	0.4736	0.2427	0.2738	0.3137	0.639	0.6919	0.7486	
<i>Volcker-Greenspan Period</i>										
<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>	
<i>2-steps ahead</i>										
<i>intercept</i>	-0.0201	0.2829	-0.0067	-0.556	-0.4786	-0.3931	0.3465	0.4354	0.5371	
<i>variance</i>	1.338	0.2114	1.3161	0.9824	1.0276	1.0798	1.6191	1.7157	1.8195	
<i>1st lag</i>	0.803	0.0435	0.8021	0.7179	0.7305	0.7481	0.858	0.8748	0.8883	
<i>inflation</i>	0.3649	0.085	0.3698	0.1767	0.221	0.2602	0.4681	0.4991	0.5335	
<i>output gap</i>	0.1422	0.0914	0.1419	-0.0361	-0.0094	0.0208	0.264	0.297	0.3291	
<i>1-step ahead</i>										
<i>intercept</i>	0.0486	0.271	0.063	-0.4593	-0.4037	-0.3106	0.3954	0.4901	0.5768	
<i>variance</i>	1.3476	0.2123	1.3252	0.9863	1.0262	1.0867	1.6331	1.7336	1.8252	
<i>1st lag</i>	0.809	0.0435	0.8093	0.7211	0.7343	0.7528	0.8624	0.8786	0.8908	
<i>inflation</i>	0.3245	0.0777	0.3271	0.1593	0.1922	0.2266	0.4205	0.4517	0.4785	
<i>output gap</i>	0.1714	0.0918	0.1721	-0.0088	0.0184	0.051	0.2898	0.3261	0.3604	
<i>0-steps ahead</i>										
<i>intercept</i>	0.1416	0.2544	0.1548	-0.3444	-0.2749	-0.1907	0.4715	0.5412	0.6403	
<i>variance</i>	1.2766	0.2026	1.2505	0.9322	0.9828	1.0283	1.5496	1.6379	1.7368	
<i>1st lag</i>	0.7931	0.0419	0.793	0.7085	0.7238	0.7406	0.8458	0.8598	0.8717	
<i>inflation</i>	0.3322	0.0683	0.3354	0.1857	0.2174	0.2442	0.4167	0.4418	0.4621	
<i>output gap</i>	0.1727	0.089	0.1698	0.0009	0.0236	0.0569	0.2868	0.3209	0.3485	

Note: Alternative Forecast Horizons - Benchmark specification with 1 lag of the dependent variable.
The last 7 columns are quantiles of the posterior distribution.

Table 5

<i>Pre-Volcker Period</i>									
<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>
<i>1:1000 iterations</i>									
<i>intercept</i>	0.9182	0.4186	0.9183	0.0999	0.2518	0.3976	1.4528	1.6212	1.8183
<i>variance</i>	0.9214	0.1931	0.8997	0.6108	0.6529	0.6972	1.1702	1.2692	1.3756
<i>1st lag</i>	0.8894	0.0453	0.8904	0.7933	0.8118	0.8284	0.9462	0.96	0.9798
<i>inflation</i>	0.0373	0.0714	0.0355	-0.097	-0.0724	-0.0489	0.1285	0.1556	0.1818
<i>output gap</i>	0.2826	0.0891	0.2831	0.1132	0.1365	0.1696	0.3957	0.4266	0.4562
<i>501:1000 iterations</i>									
<i>intercept</i>	0.9251	0.4275	0.9143	0.0822	0.2461	0.3976	1.4669	1.6311	1.8212
<i>variance</i>	0.9234	0.1928	0.8963	0.6113	0.65	0.6997	1.1844	1.2682	1.3756
<i>1st lag</i>	0.8892	0.0451	0.8899	0.7914	0.8118	0.8268	0.9463	0.9603	0.9792
<i>inflation</i>	0.0361	0.0698	0.0345	-0.0947	-0.0695	-0.0451	0.1279	0.1556	0.1894
<i>output gap</i>	0.2827	0.0882	0.2811	0.1134	0.1404	0.1686	0.3974	0.4259	0.4575
<i>1:500 iterations</i>									
<i>intercept</i>	0.9114	0.4097	0.9225	0.1075	0.2518	0.392	1.4289	1.5814	1.7918
<i>variance</i>	0.9193	0.1936	0.9026	0.6108	0.6529	0.6933	1.1533	1.2692	1.3774
<i>1st lag</i>	0.8895	0.0455	0.8906	0.7983	0.8107	0.8294	0.9459	0.9571	0.9805
<i>inflation</i>	0.0385	0.0731	0.0375	-0.097	-0.0748	-0.0534	0.1285	0.1556	0.1818
<i>output gap</i>	0.2825	0.0901	0.2837	0.1132	0.1342	0.17	0.3939	0.4283	0.4562
<i>Volcker-Greenspan Period</i>									
<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>
<i>1:1000 iterations</i>									
<i>intercept</i>	-0.0541	0.2651	-0.0426	-0.5502	-0.4878	-0.3975	0.2937	0.3796	0.4612
<i>variance</i>	1.1051	0.179	1.0882	0.7965	0.8413	0.8926	1.3393	1.425	1.5113
<i>1st lag</i>	0.7823	0.0407	0.7814	0.6999	0.7167	0.7292	0.8331	0.8482	0.8575
<i>inflation</i>	0.428	0.0806	0.4296	0.2529	0.2861	0.3301	0.5257	0.5528	0.5884
<i>output gap</i>	0.1112	0.0825	0.11	-0.0495	-0.0267	0.0044	0.2184	0.2524	0.2785
<i>501:1000 iterations</i>									
<i>intercept</i>	-0.0495	0.2557	-0.0418	-0.5278	-0.4698	-0.3801	0.2871	0.3599	0.4514
<i>variance</i>	1.0899	0.1677	1.0708	0.7927	0.8348	0.897	1.3152	1.3709	1.4946
<i>1st lag</i>	0.7834	0.0394	0.7818	0.7069	0.7181	0.7314	0.8337	0.8507	0.8573
<i>inflation</i>	0.4248	0.0783	0.4275	0.2527	0.2893	0.3225	0.5221	0.5447	0.5715
<i>output gap</i>	0.1086	0.079	0.1101	-0.0494	-0.0274	0.0057	0.2094	0.2498	0.2824
<i>1:500 iterations</i>									
<i>intercept</i>	-0.0587	0.2743	-0.0447	-0.5719	-0.5089	-0.4294	0.3076	0.3849	0.4683
<i>variance</i>	1.1204	0.1885	1.1096	0.8113	0.8432	0.8896	1.3646	1.4557	1.5208
<i>1st lag</i>	0.7812	0.042	0.7801	0.6946	0.7105	0.7254	0.8326	0.8469	0.8628
<i>inflation</i>	0.4312	0.0828	0.4319	0.2571	0.2858	0.3318	0.5274	0.564	0.6088
<i>output gap</i>	0.1137	0.0859	0.1094	-0.0495	-0.0241	0.0022	0.2264	0.2541	0.2763

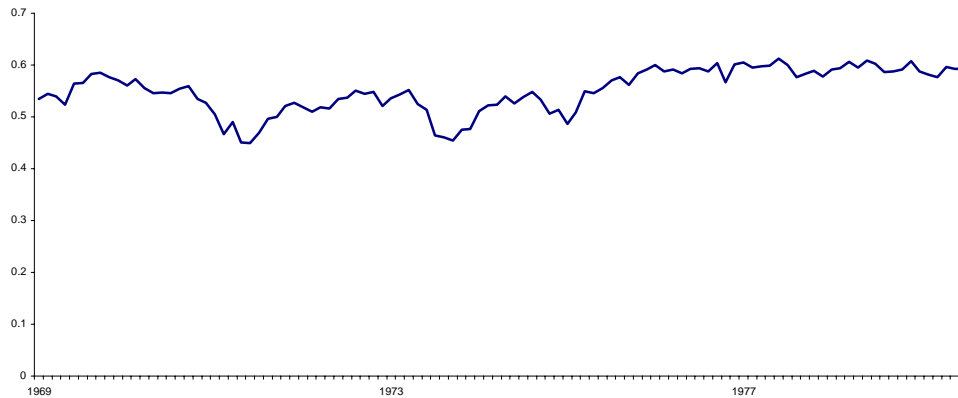
Note: Specification with alternative threshold values.
 The last 7 columns are quantiles of the posterior distribution.

Table 6

2-state Markov Switching specification for variance

<i>variable</i>	<i>mean</i>	<i>std. dev.</i>	<i>median</i>	<i>2.5% qntl</i>	<i>5% qntl</i>	<i>10% qntl</i>	<i>90% qntl</i>	<i>95% qntl</i>	<i>97.5% qntl</i>
<i>Pre-Volcker Period</i>									
<i>intercept</i>	0.99	0.5052	0.9887	0.0204	0.1702	0.3685	1.6201	1.8504	1.9956
<i>low variance</i>	1.0908	0.2325	1.0653	0.7027	0.7563	0.8163	1.3997	1.5395	1.6322
<i>high variance</i>	1.5684	1.0332	1.3651	0.8326	0.8968	0.9733	2.1263	2.7755	3.4461
<i>1st lag</i>	0.709	0.1266	0.713	0.464	0.4993	0.5416	0.8618	0.916	0.9547
<i>2nd lag</i>	0.1737	0.1226	0.1777	-0.0712	-0.021	0.0194	0.3296	0.3742	0.4084
<i>inflation</i>	0.0448	0.0826	0.0437	-0.1146	-0.091	-0.0614	0.148	0.1816	0.2078
<i>output gap</i>	0.3325	0.1107	0.3321	0.1091	0.1441	0.193	0.4743	0.5104	0.5524
<i>Volcker-Greenspan Period</i>									
<i>intercept</i>	-0.1824	0.3458	-0.1776	-0.8788	-0.7454	-0.6027	0.2291	0.3685	0.5451
<i>low variance</i>	1.4258	0.239	1.4132	0.9996	1.0579	1.1272	1.7444	1.8517	1.9457
<i>high variance</i>	11.0306	14.942	6.8423	1.3883	1.6035	2.7019	22.4079	33.1521	42.0234
<i>1st lag</i>	0.6465	0.1217	0.6513	0.4169	0.4351	0.479	0.7984	0.8337	0.8796
<i>2nd lag</i>	0.1116	0.112	0.1083	-0.1085	-0.0786	-0.0262	0.2675	0.2992	0.327
<i>inflation</i>	0.5182	0.1139	0.5209	0.2981	0.3237	0.3698	0.6639	0.6951	0.7363
<i>output gap</i>	0.0998	0.1016	0.0988	-0.109	-0.0657	-0.0226	0.23	0.2688	0.3052

Pre-Volcker low state probabilities



Volcker-Greenspan low state probabilities

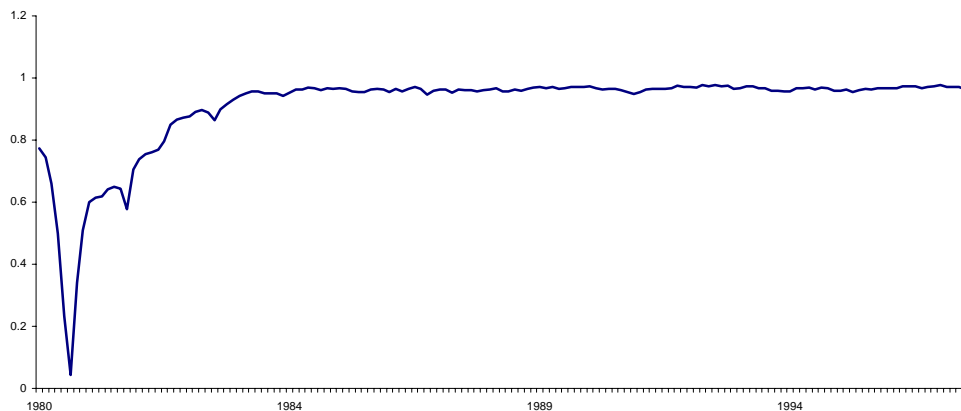


Figure 4A: Time Varying Parameters (and 90% bands), Pre-Volcker Period

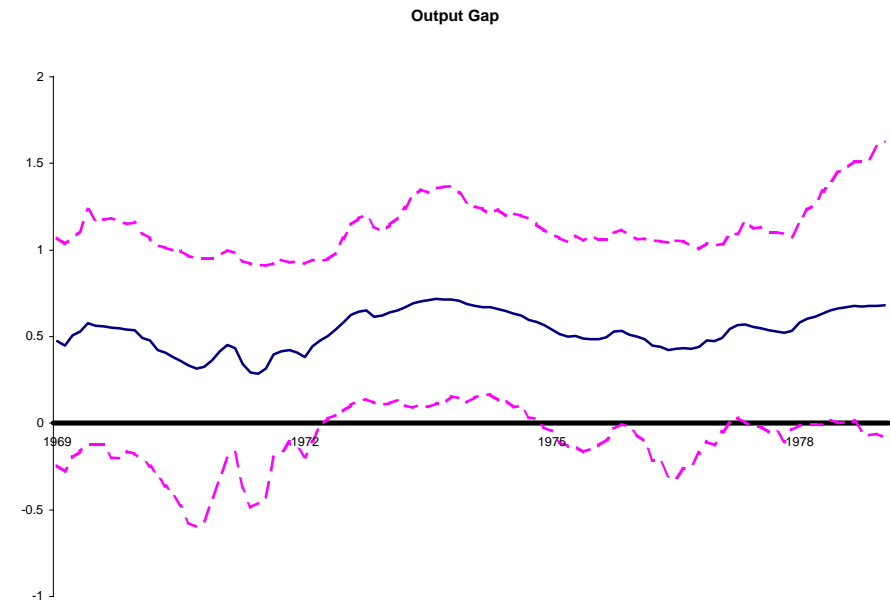
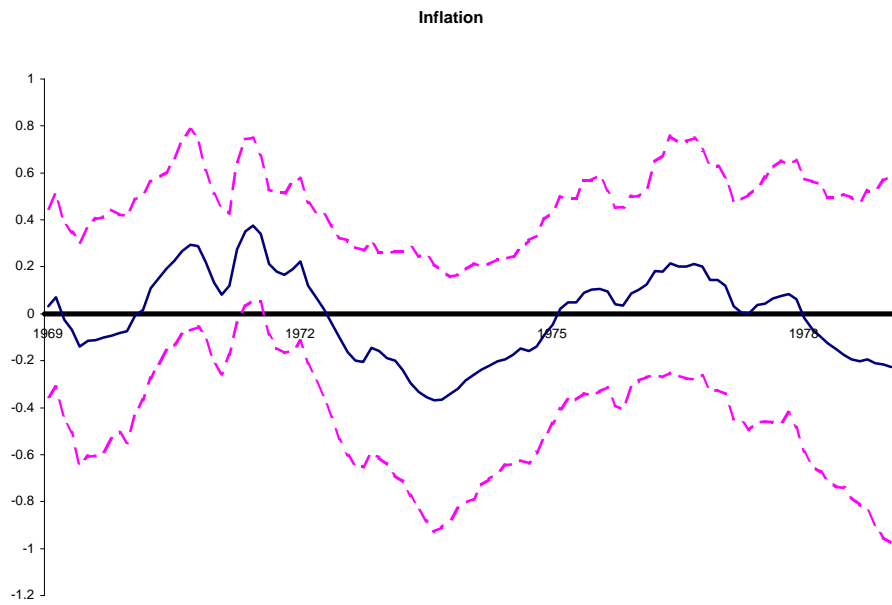
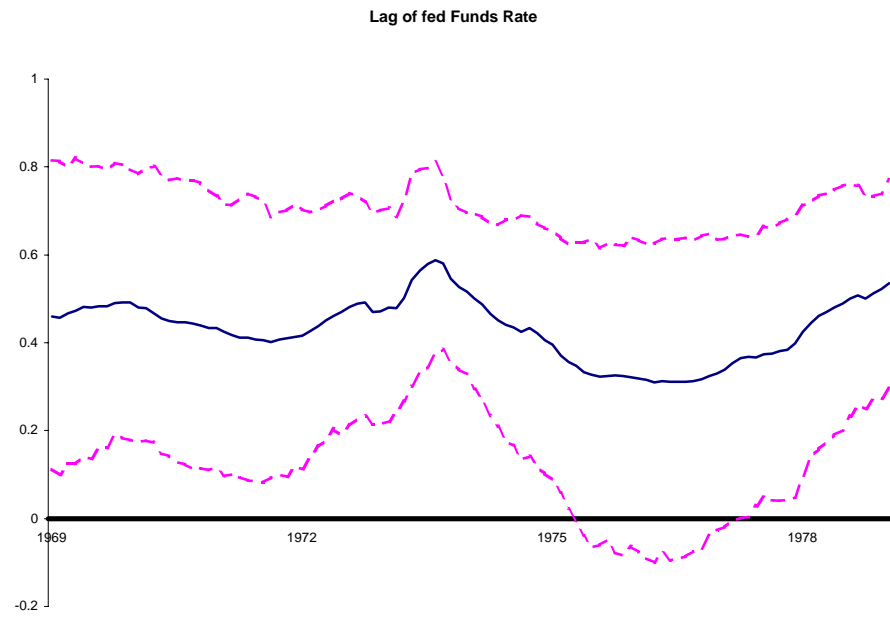
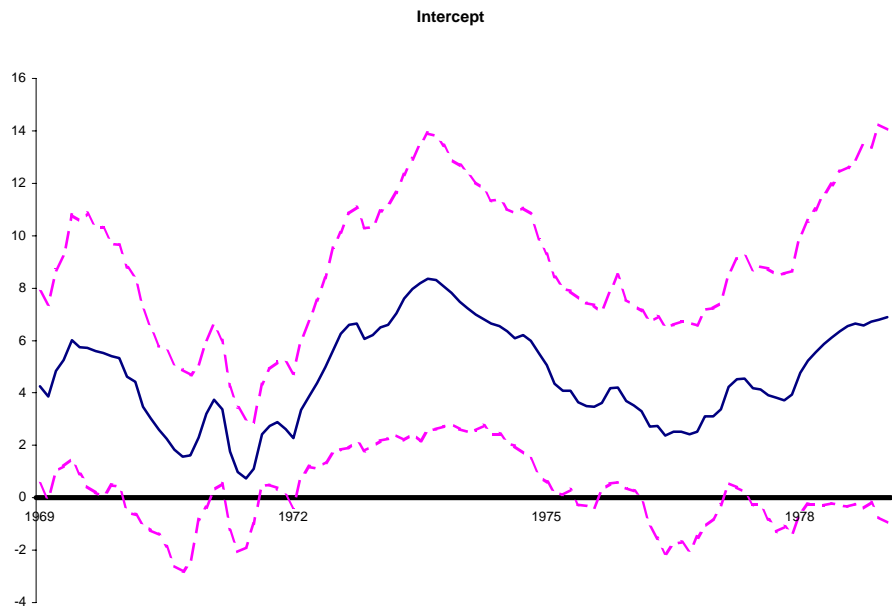
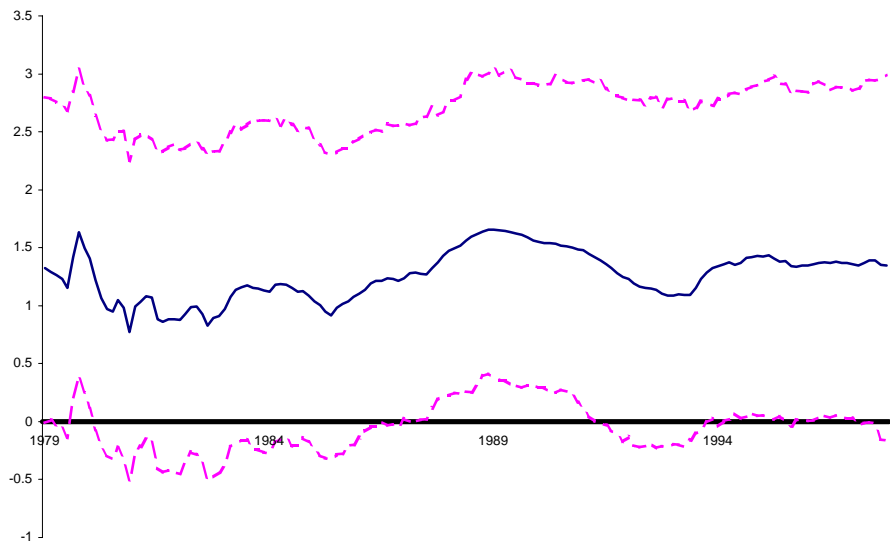
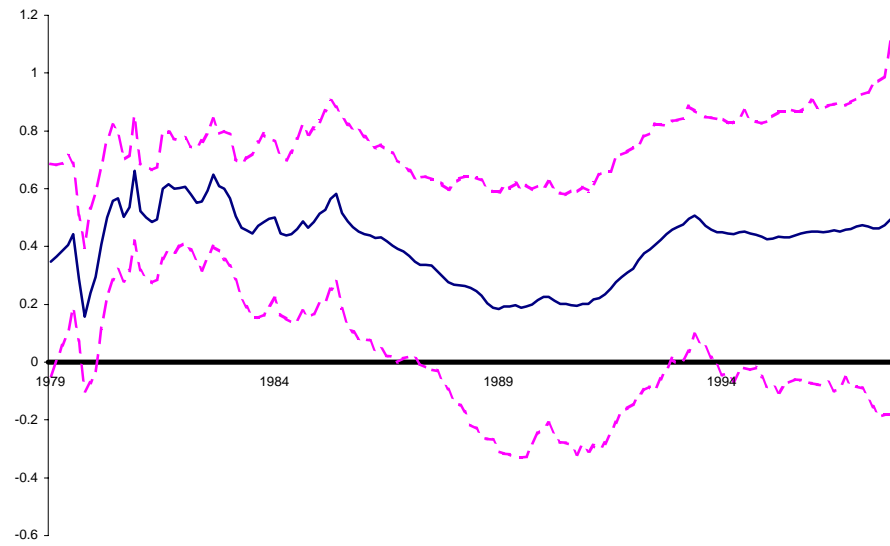


Figure 4B: Time Varying Parameters (and 90% bands), Volcker-Greenspan Period

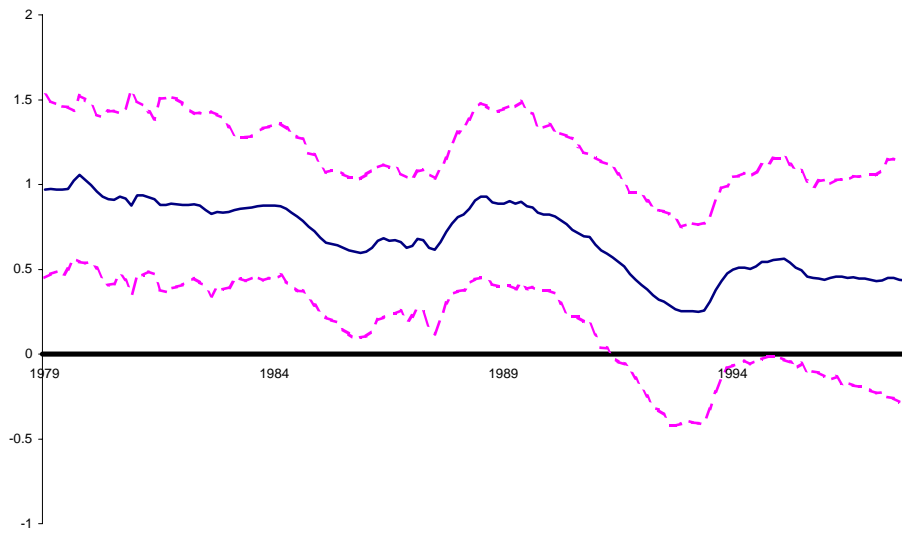
Intercept



Lag of Fed Funds Rate



Inflation



Output Gap

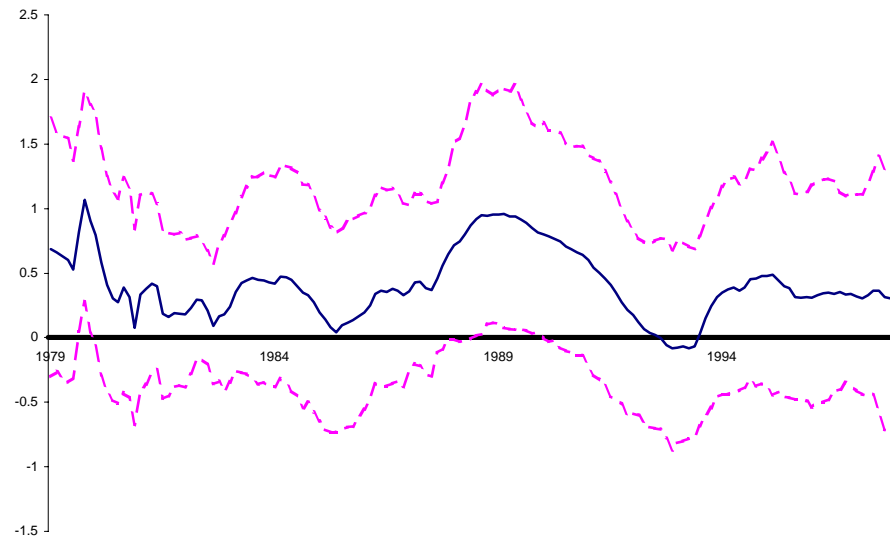
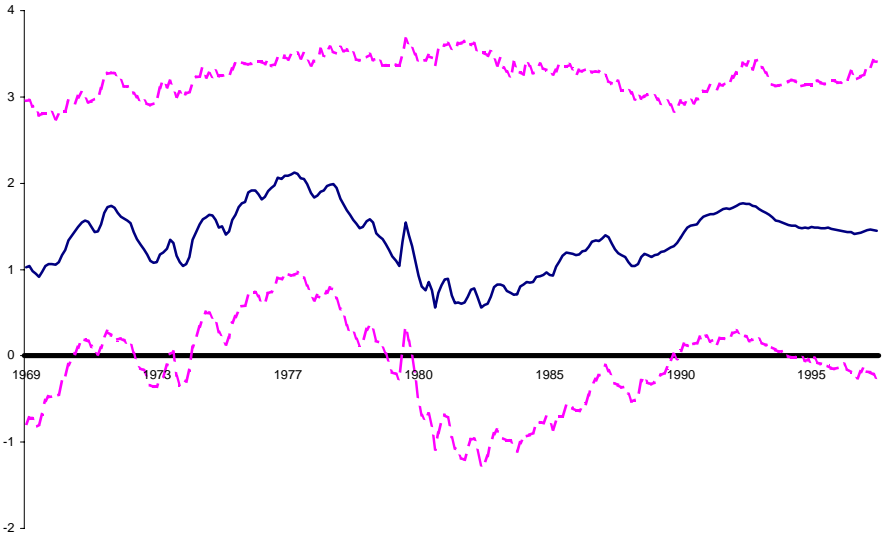
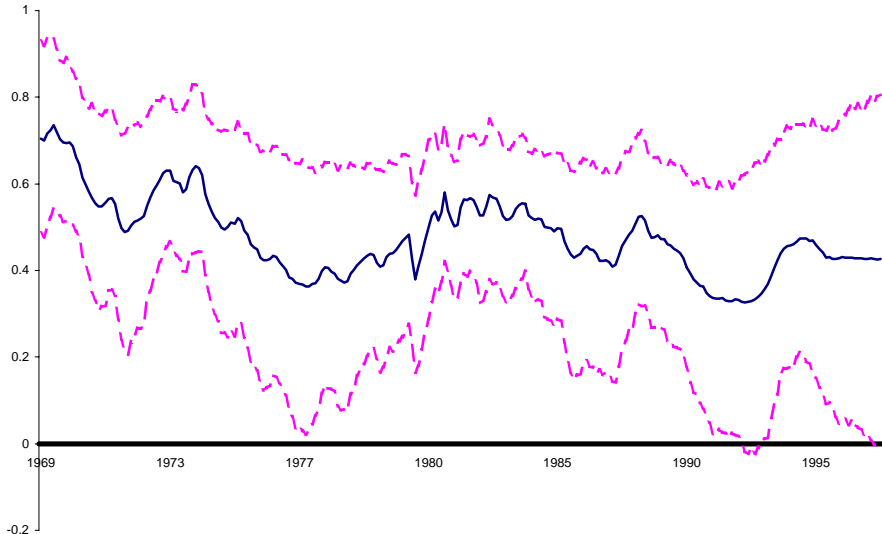


Figure 4C: Time Varying Parameters (and 90% bands), 1969-1998

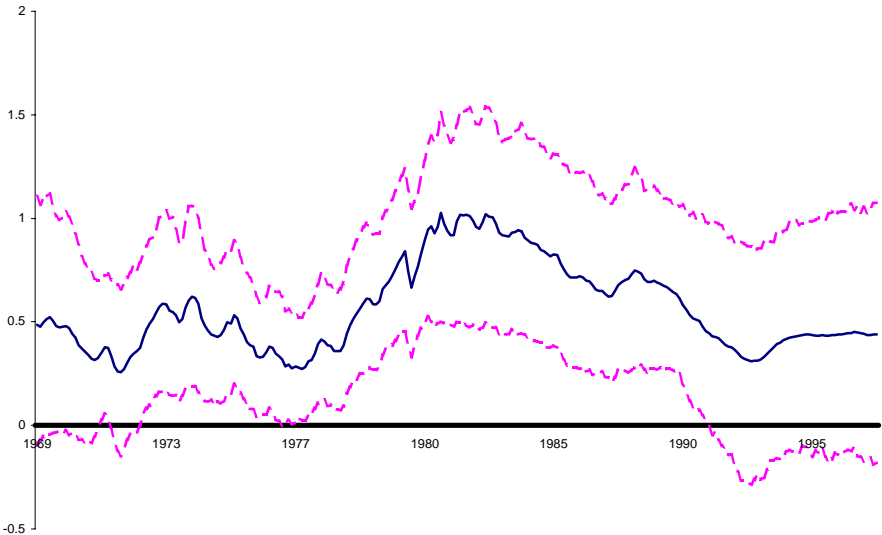
Intercept



Lag of Fed Funds Rate



Inflation



Output Gap

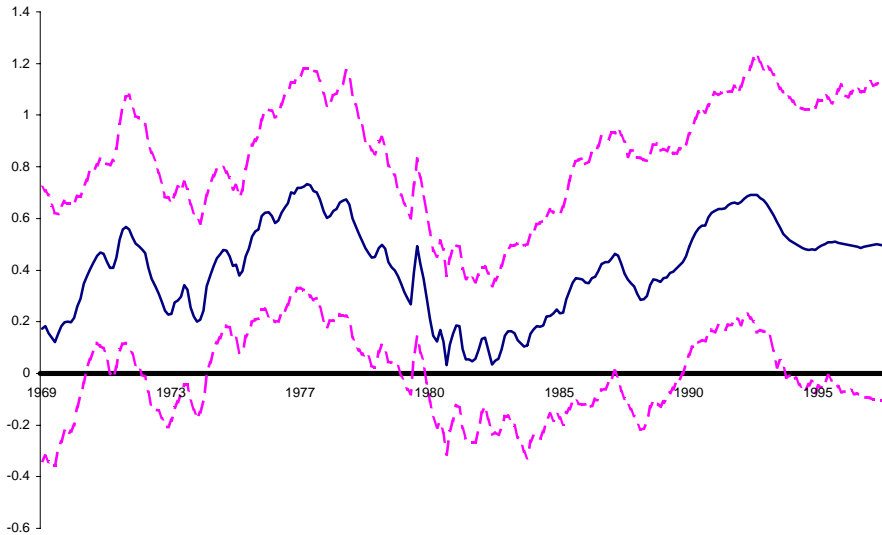
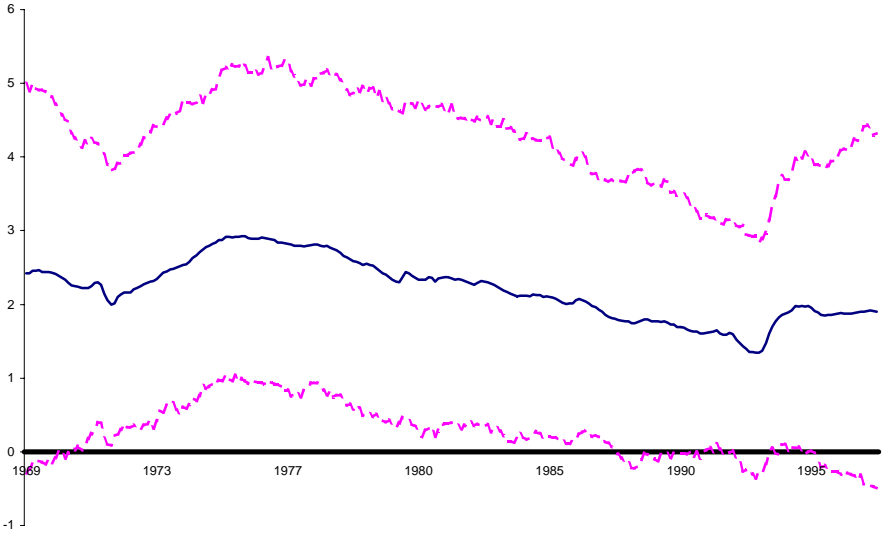
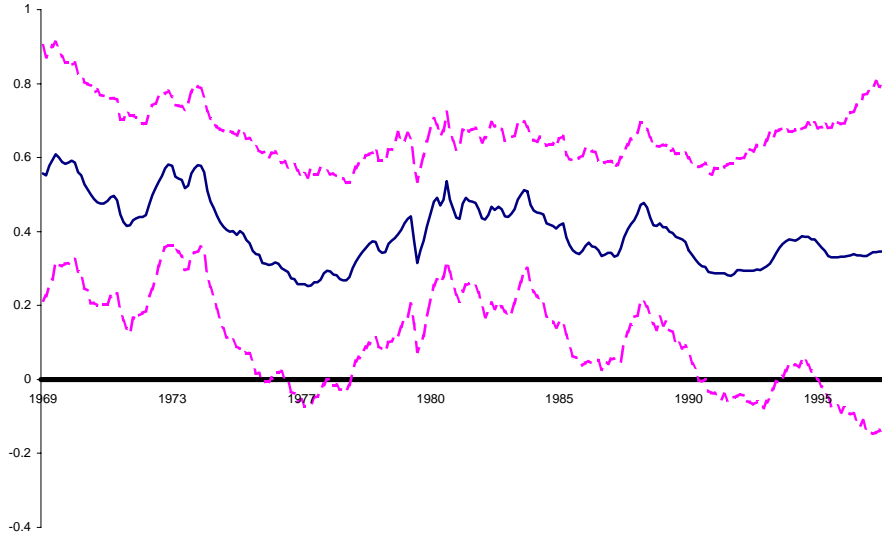


Figure 4D: Time Varying Parameters (and 90% bands), 1969-1998
 (allowing for the possibility of a jump in August of 1979)

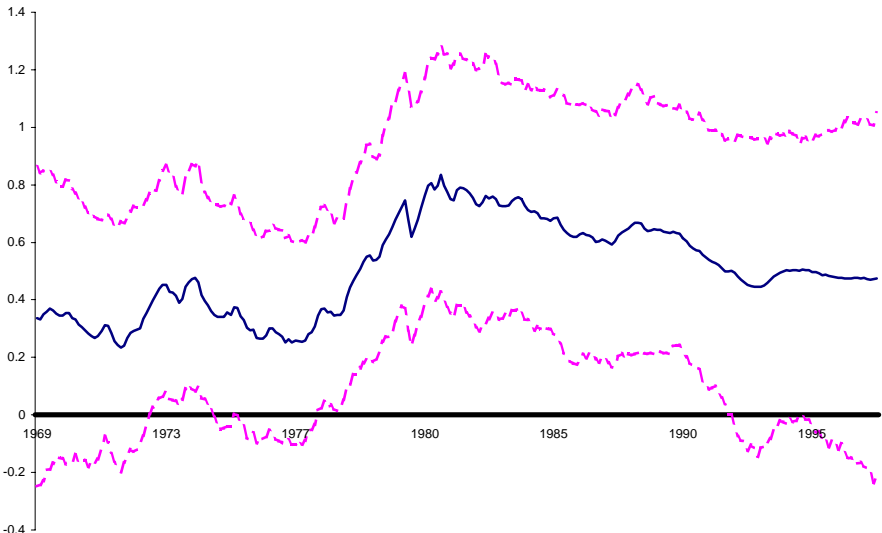
Intercept



Lag of Fed Funds Rate



Inflation



Output Gap

