Evolving Post-World War II U.S. Inflation Dynamics*

Timothy Cogley
Arizona State University

Thomas J. Sargent
Stanford University and Hoover Institution

June 2001

Abstract

For postwar U.S. data, this paper uses Bayesian methods to account for the four sources of uncertainty in a random coefficients VAR for inflation, unemployment, and an interest rate. We use the model to assemble evidence about the evolution of measures of the persistence of inflation, prospective long-horizon forecasts (means) of inflation and unemployment, statistics for testing an approximation to the natural unemployment rate hypothesis, and a version of a Taylor rule. We relate these measures to stories that interpret the conquest of U.S. inflation under Volcker and Greenspan as reflecting how the monetary policy authority came to learn an approximate version of the natural unemployment rate hypothesis. We study Taylor's warning that defects in that approximation may cause the monetary authority to forget the natural rate hypothesis as the persistence of inflation attentuates.

*We are grateful to Irena Asmundsen, Sergei Morozov, and Chao Wei for excellent research assistance. For comments and suggestions, we thank Charles Evans, Marvin Goodfriend, Lars Hansen, Chang-Jin Kim, Robert King, Charles Nelson, Simon Potter, Martin Schneider, Christopher Sims, James Stock, Harald Uhlig, and seminar participants at ASU, FRB Atlanta, FRB Richmond, Penn, Stanford, UCLA, UC Riverside, UC Santa Barbara, the 2000 SED Meetings, and the 2001 NBER Macro Annual Conference. Sargent thanks the National Science Foundation for research support through a grant to the National Bureau of Economic Research.

1. Introduction

This paper uses a nonlinear stochastic model to describe inflation-unemployment dynamics in the U.S. after World War II. The model is a vector autoregression with coefficients that are random walks with reflecting barriers that keep the VAR stable. The innovations in the coefficients are arbitrarily correlated with each other and with innovations to the observables. The model enables us to detect features that have been emphasized in theoretical analyses of inflation-unemployment dynamics. Those analyses involve coefficient drift in essential ways.

Thus, DeLong (1997), Taylor (1997, 1998), and Sargent (1999) interpreted the broad movements of the inflation rate in terms of the monetary authority's changing views about the Phillips curve. According to them, the run-up in inflation in the late 1960's and 1970's occurred because the monetary authority believed that there was an exploitable trade-off between inflation and unemployment. Its beliefs induced the monetary authority to accept the temptation to inflate more and more until eventually it had attained Kydland-Prescott (1977) time-consistent inflation rates. But the observations of the 1970's taught Volcker and Greenspan the natural rate hypothesis, which they eventually acted upon to reduce inflation.

Another mechanism was posited by Parkin (1993) and Ireland (1999), who argued that the inflation-unemployment dynamics are driven by exogenous drift in the natural rate of unemployment, for example due to demographic changes. Because the time-consistent inflation rate varies directly with the natural rate of unemployment, Parkin and Ireland attributed the drift in the inflation rate to drift in the natural rate of unemployment.

The DeLong-Taylor-Sargent story makes contact with various elements in Lucas's (1976) Critique. It makes the drift in inflation-unemployment dynamics a consequence of the monetary authority's evolving views about the economy. The story attributes alterations in the law of motion for inflation and unemployment to the changing behavior of the monetary authority, which emerges in turn from its changing beliefs. This story is consistent with one way that Lucas (1976) has been read, namely, as an invitation to impute observed drift in coefficients of econometric models to time-series variation in government policy functions.

Sargent's (1999) version of the story focuses on how the coefficient drift over time affected the results of time-series tests of the natural rate hypothesis. In the late 1960's, Robert Solow and James Tobin proposed a test of the natural rate hypothesis. Using data through the late 1960's that test rejected the natural rate hypothesis in favor of a permanent trade-off between inflation and unemployment. Lucas (1972) and Sargent (1971) criticized that test for not properly stating the implications of the natural rate hypothesis under rational expectations. In particular, the Solow-Tobin test was correct only if inflation exhibited a unit root. Before the 1970's, post-war U.S. inflation data did not exhibit a unit root, rendering invalid (in the opinion of Lucas and Sargent) Solow's and Tobin's interpretation of their test. However, in the 1970's, just when U.S. inflation seems to have acquired a unit root, the Solow-Tobin test began accepting the natural rate hypothesis. Building on Sims (1988) and Chung (1990), Sargent (1999) constructs an adaptive model of the government's learning and policy making that centers on the process by which the government learns an imperfect version of the natural rate hypothesis, cast in terms of Solow and Tobin's representation.

Parts of Sargent's adaptive story acquire credibility when it is noted how the Solow-Tobin characterization of the natural rate hypothesis has endured, despite the criticism of Lucas and Sargent. As Hall (1999) and Taylor (1998) lament, that faulty characterization continues to be widely used. For example, see Rudebusch and Svensson (1999) for a widely cited model that represents the natural rate hypothesis in the Solow-Tobin form. Fisher and Seater (1993), King and Watson (1994 and 1997), Fair (1996), Eisner (1997), and Ahmed and Rogers (1998) construct tests of long-run neutrality that are predicated on the assumption of a unit root in inflation.[1] Estrella and Mishkin (1999) use the Solow-Tobin characterization to estimate the natural rate of unemployment. In the discussion following the paper by Estrella and Mishkin, John Williams confesses that the Federal Reserve Board's large-scale macroeconometric model also incorporates this characterization. Hall questions its validity for U.S. data after 1979 and sharply criticizes its continued use.

Taylor (1998) warns that adherence to the erroneous econometric characterization of the natural rate hypothesis will eventually cause policy to go astray. Because of the diminished serial correlation that he sees in recent inflation data, Taylor is concerned that the disappearance of a unit root in inflation means that the faulty test may soon signal an exploitable trade-off that will once again tempt the monetary authority. The theme of both Hall and Taylor is that failure to remember the theoretical and econometric lessons of the 1970's is likely to resuscitate pressure to inflate emanating from the empirical Phillips curve. In the same

[1]Many of these authors pre-test for a unit root and apply the Solow-Tobin test only if they fail to reject the null hypothesis. But pretesting could result in a more subtle version of the Lucas-Sargent trap. Unit root tests have low power and may fail to detect circumstances in which the Solow-Tobin test is inappropriate.

symposium, Friedman (1998) and Solow (1998) made a number of assertions that may have contributed to Taylor's worries. Friedman asserted that the real effects of monetary policy are so long-lasting that "for all practical purposes they might just as well be permanent." Solow (1998) expressed skepticism about the natural rate hypothesis and suggested that the supporting evidence is specific to the U.S. economy since 1970. He argued that monetary policy can affect the natural rate of unemployment and that the experience of the U.S. in the 1960's suggests that persistent high unemployment would yield to a revival of aggregate demand. Taylor's concern is that low inflation would be hard to sustain if belief in a long-run trade-off were again to become influential.

The object of this paper is to develop empirical evidence that is relevant to this discussion.[2] Section 2 describes a Bayesian model that we use to summarize the evolution of inflation dynamics. Section 3 reports stylized facts about this evolution, and section 4 discusses test statistics for the Solow-Tobin version of the natural rate hypothesis. Section 5 considers Taylor's warning about recidivism on the natural rate hypothesis. The paper concludes with a summary.

2. A Random Coefficients Representation

We use a Bayesian vector autoregression with time-varying parameters to describe the evolution of the law of motion for inflation. We are interested in a random coefficients representation for some of the reasons expressed in the initial sections of Lucas (1976). The Bayesian framework treats coefficients as random variables, making it attractive for modeling data from economies in which important decision makers, including the monetary authority, are learning.[3]

---

[2] Albanesi, Chari, and Christiano (2000) model the inception and termination of inflation in the 1970's with a sunspot variable that shifts expectations between two regimes. Their equilibrium excludes the concerns about model misspecification that are the focus of the present discussion. It is possible that a regime switching model like theirs can confront the observations about co-movements between inflation persistence and mean inflation that we document below.

[3] Our focus in this paper is on the evolution of reduced form relationships. Structural models involve nonlinear cross-equation restrictions on the evolving parameters, and they require nonlinear filtering methods. We are currently studying nonlinear filters.

2.1. Notation and State-Space Representation

The model has a non-linear state-space representation. The measurement equation is

$$y_t = X_t'\theta_t + \varepsilon_t, \tag{2.1}$$

where y_t is an $N \times 1$ vector of endogenous variables, θ_t is a $K \times 1$ vector of coefficients, X_t' is an $N \times K$ matrix of predetermined and/or exogenous variables, and ε_t is an $N \times 1$ vector of prediction errors. The vector y_t includes inflation and variables useful for predicting inflation. In this paper, we use (2.1) to represent a vector autoregression, so that the right-hand variables are lags of y_t. In an unrestricted vector autoregression, each equation contains the same right-hand variables, $X_t' = (I_N \otimes x_t')$.

We treat the coefficients of the VAR as a hidden state vector. The state vector θ_t evolves according to

$$p(\theta_{t+1}|\theta_t, V) \propto I(\theta_{t+1}) f(\theta_{t+1}|\theta_t, V) \tag{2.2}$$

where $I(\theta_t) = 0$ if the roots of the associated VAR polynomial are inside the unit circle and 1 otherwise; V is a covariance matrix defined below; and

$$f(\theta_{t+1}|\theta_t, V) \sim N(\theta_t, Q). \tag{2.3}$$

Thus, $f(\theta_{t+1}|\theta_t, V)$ can be represented as the driftless random walk

$$\theta_t = \theta_{t-1} + v_t, \tag{2.4}$$

where v_t is an i.i.d. Gaussian process with mean 0 and covariance Q. The economy changes over time when news arrives, making θ_t vary in an unpredictable way. Throughout this paper, we use $f(\cdot)$ to denote a normal density, and $p(\cdot)$ to denote a more general density.

We assume that the innovations, $(\varepsilon_t', v_t')'$, are identically and independently distributed normal random variables with mean zero and covariance matrix,

$$E_t \begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \begin{bmatrix} \varepsilon_t' & v_t' \end{bmatrix} = V = \begin{pmatrix} R & C' \\ C & Q \end{pmatrix}, \tag{2.5}$$

where R is the $N \times N$ covariance matrix for measurement innovations, Q is the $K \times K$ covariance matrix for state innovations, and C is a $K \times N$ cross-covariance matrix. Following the Bayesian literature, we call the θ's "parameters" and the elements of R, Q, and C "hyperparameters."

We assume that the hyperparameters and initial state θ_0 are independent, that the initial state is a truncated Gaussian random variable, and that the hyperparameters come from an inverse-Wishart distribution. We adopted these parts of the prior mostly because of their convenience in being natural conjugates for our Gaussian virtual prior f.

Let $f(\theta_0) = N(\bar{\theta}, \bar{P})$ represent a normal prior with mean $\bar{\theta}$ and variance $\bar{P}$. The prior for the initial state is

$$p(\theta_0) \propto I(\theta) N(\bar{\theta}, \bar{P}). \tag{2.6}$$

Our prior for the hyperparameters is

$$p(V) = IW(\overline{V}^{-1}, T_0), \tag{2.7}$$

where $IW(S, df)$ represents the inverse-Wishart distribution with scale matrix S and degrees of freedom df. This is a convenient form because it yields an inverse-Wishart posterior when combined with a Gaussian likelihood. Collecting the pieces, the joint prior for θ_0, V can be represented as

$$p(\theta_0, V) \propto I(\theta) N(\bar{\theta}, \bar{P}) IW(\overline{V}^{-1}, T_0). \tag{2.8}$$

Both pieces are informative, but in the empirical section we set $\bar{\theta}$, $\bar{P}$, $\overline{V}$, and T_0 so that they are only weakly informative.

We use the following notation to denote partial histories of the variables Y_t and θ_t. The vectors

$$Y^T = [y_1', ..., y_T']' \tag{2.9}$$

and

$$\theta^T = [\theta_1', ..., \theta_T']' \tag{2.10}$$

represent the history of data and states up to date T and

$$Y^{T+1,T+H} = [y_{T+1}', ..., y_{T+H}']' \tag{2.11}$$

and

$$\theta^{T+1,T+H} = [\theta_{T+1}', ..., \theta_{T+H}']' \tag{2.12}$$

represent potential future trajectories from date T onward.

We can use (2.2) to assemble the joint density

$$p(\theta^T|V) \propto I(\theta^T)f(\theta^T|V) \tag{2.13}$$

where

$$f(\theta^T|V) = f(\theta_0|V)\prod_{t=0}^{T-1} f(\theta_{t+1}|\theta_t, V) \tag{2.14}$$

and

$$I(\theta^T) = \prod_{t=0}^{T} I(\theta_t). \tag{2.15}$$

We call f our 'virtual prior' and p the prior. The virtual prior f makes θ a driftless random walk. Multiplying $f(\theta^T|V)$ by $I(\theta^T)$ puts zero probability on sample paths of $\{\theta_t\}$ for which θ_t for any $t \geq 0$ corresponds to unstable VAR coefficients.[4]

In (2.2), the truncation of $f(\theta_t|\theta_{t-1}, V)$ through multiplication by $I(\theta_t)$ reflects our opinion that explosive representations are implausible for the United States. An unrestricted normal density $f(\theta^T|V) = f(\theta_0)\prod_{t=0}^{T} f(\theta_{t+1}|\theta_t, V)$ for the history of states θ^T implies a positive probability of explosive autoregressive roots, but an explosive representation implies an infinite variance for inflation, which cannot be optimal for a central bank that minimizes a loss function involving the variance of inflation.[5] We restrict the prior to put zero probability on explosive states.

This representation resembles some of the models in Doan, Litterman, and Sims (1984), but with a different prior. Doan, et. al. were primarily interested in forecasting and recommended a "random walk in variables" prior to promote parsimony. We are less interested in forecasting and more interested in summarizing the data in a relatively unconstrained fashion, so we chose the prior described above.

[4]An appendix shows that the model formed by (2.3), (2.13), (2.14), and (2.15) implies the nonlinear transition equation (2.2).

[5]Alternatively, explosive representations cannot result if the monetary policy rule ensures that inflation is bounded. We do not claim that an integrated representation for inflation is implausible on statistical grounds, only that drift in inflation is hard to reconcile with purposeful central bank behavior.

2.2. A Limitation of Our Model: No Stochastic Volatility

For macroeconomic variables and a period similar to ours, Bernanke and Mihov (1998a, 1998b) and Sims (1999) presented evidence that favors a vector autoregression with time-invariant autoregressive coefficients but a covariance matrix of innovations that fluctuates over time. In contrast, our specification allows the coefficients to vary and assumes a time-invariant but unknown innovation covariance matrix V. While our prior fixes V, our statistical methods nevertheless allow the data to speak up for volatility or drift in V, albeit in a restricted and adaptive way. Our estimates of V conditioned on time t data fluctuate over time in ways that we shall discuss.

We chose our specification partly because we want to focus attention on the coefficient drift issues raised by Lucas (1976). Our model is rigged to let us detect drifts in the systematic parts of government and private behavior rules that show up in the systematic parts of vector autoregressions. Our prior embodies a prejudice that monetary policy changed systematically during the years that we study. In contradistinction, the interpretation of the evidence favored by Bernanke and Mihov (1998a, 1998b) and Sims (1999) is consistent with a view that while distributions of shocks have evolved, agents' responses to them have been stable.[6]

2.3. Posterior Predictive Density

As Bayesians, our goal is to summarize the posterior density for the objects of interest. We are mostly interested in a forward looking perspective on inflation, so we want posterior *predictive* densities.

In this model, there are four sources of uncertainty about the future. The terminal state θ_T and the hyperparameters V are unknown and must be estimated. In addition, as time goes forward the state vector will drift away from θ_T, and the measurement equation will be hit by random shocks. Conditional on prior beliefs and data through date T, beliefs about the future can be expressed by the joint posterior distribution,

$$p(Y^{T+1,T+H}, \theta^{T+1,T+H}, \theta^T, V \,|Y^T). \tag{2.16}$$

Our objective is to characterize (2.16). This is a complicated object, but it can be decomposed into more tractable components. We begin by factoring (2.16)

[6] See Sims (1982) and Sargent (1983) for theoretical settings that, by assuming that the historical sample was produced by optimizing government behavior and stable private sector responses to it, can explain such a pattern.

into the product of a conditional and a marginal density,

$$p(Y^{T+1,T+H}, \theta^{T+1,T+H}, \theta^T, V\,|Y^T) = p(\theta^T, V\,|\,Y^T) \times \\ p(Y^{T+1,T+H}, \theta^{T+1,T+H}\,|\theta^T, V, Y^T). \tag{2.17}$$

This expression splits the joint density into a term that represents beliefs about the past and present and another that represents beliefs about the future. The first term is the joint posterior density for hyperparameters and the history of states. It summarizes current knowledge about system dynamics, based on data and prior beliefs. The second term reflects the uncertainty about the future that would be present even if the current state and hyperparameters were known with certainty. This term reflects the influence of future innovations to the state and measurement equations.

Analytical expressions for each piece are unavailable, even for simple cases. Instead, we use Monte Carlo methods to simulate them. The algorithm is split into two parts, corresponding to the components of (2.17). The first part uses the Gibbs sampler to simulate a draw of θ^T and V from the marginal density, $p(\theta^T, V\,|\,Y^T)$. The second step plugs that draw into the conditional density $p(Y^{T+1,T+H}, \theta^{T+1,T+H}\,|\theta^T, V, Y^T)$ and generates a trajectory for future data and states.

2.4. Beliefs About the Past and Present

The posterior density for states and hyperparameters can be expressed as

$$\begin{aligned} p(\theta^T, V\,|\,Y^T) &\propto p(Y^T\,|\,\theta^T, V)p(\theta^T, V), \\ &\propto f(Y^T\,|\,\theta^T, V)p(\theta^T|V)p(V), \\ &\propto I(\theta^T)\left[f(Y^T\,|\,\theta^T, V)f(\theta^T|V)p(V)\right]. \end{aligned} \tag{2.18}$$

The first line follows from Bayes' theorem: $p(\theta^T, V)$ represents a joint prior for hyperparameters and states and $p(Y^T\,|\,\theta^T, V)$ is a conditional likelihood. Conditional on states and hyperparameters, the measurement equation is linear in observables and has normal innovations. Thus, the conditional likelihood is Gaussian, $p(Y^T\,|\,\theta^T, V) = f(Y^T\,|\,\theta^T, V)$, as shown in the second line. The joint prior for hyperparameters and states can be factored into a marginal prior for V and a conditional prior for θ^T, and substituting $I(\theta^T)f(\theta^T|V)$ for $p(\theta^T|V)$ delivers the expression on the third line.

Notice that the term in brackets on the last line is the joint posterior kernel that would result if the restriction on unstable roots were not imposed. If not for this restriction, the model would have a linear Gaussian state-space representation, with transition equation $f(\theta^T|V)$. The posterior kernel associated with this linear transition law is

$$p_L(\theta^T, V \mid Y^T) \propto f(Y^T \mid \theta^T, V) f(\theta^T|V) p(V). \tag{2.19}$$

Substituting this relation into the last equation, the posterior density for the nonlinear model can be expressed as a truncation of the posterior for the unrestricted linear model,

$$p(\theta^T, V \mid Y^T) \propto I(\theta^T) p_L(\theta^T, V \mid Y^T). \tag{2.20}$$

Among other things, this means that $p(\theta^T, V \mid Y^T)$ can be represented and simulated in two steps. First, we derive the posterior associated with linear transition equation, $p_L(\theta^T, V \mid Y^T)$, and then we multiply by $I(\theta^T)$ to rule out explosive outcomes. In the Monte Carlo simulation, this is implemented by simulating the unrestricted posterior and rejecting draws that violate the stability condition. The next subsection describes our method for simulating $p_L(\theta^T, V \mid Y^T)$, and the one after that confirms the validity of our rejection sampling procedure.

2.5. Simulating the Unrestricted Posterior

Following Kim and Nelson (1999), we use the Gibbs sampler to simulate draws from $p_L(\theta^T, V \mid Y^T)$. The Gibbs sampler iterates on two operations. First, conditional on the data and hyperparameters, we draw a history of states from $p_L(\theta^T | Y^T, V)$. Then, conditional on the data and states, we draw hyperparameters from $p_L(V \mid Y^T, \theta^T)$. Subject to regularity conditions (see Roberts and Smith 1992), the sequence of draws converges to a draw from the joint distribution, $p_L(\theta^T, V \mid Y^T)$.

2.5.1. Gibbs Step 1: States Given Hyperparameters

Conditional on data and hyperparameters, the unrestricted transition law is linear and has normal innovations. Thus, the virtual states are Gaussian,

$$p_L(\theta^T | Y^T, V) = f(\theta^T \mid Y^T, V). \tag{2.21}$$

This density can be factored as[7]

$$f(\theta^T \,|\, Y^T, V) = f(\theta_T \,|\, Y^T, V) \prod_{t=1}^{T-1} f(\theta_t \,|\, \theta_{t+1}, Y^t, V). \tag{2.22}$$

The leading term is the marginal posterior for the terminal state, and the other terms are conditional densities for the preceding time periods. Since the conditional densities on the right hand side are Gaussian, it is enough to update their conditional means and variances. This can be done via the Kalman filter.

Deriving forward and backward recursions for $f(\theta^T | Y^T, V)$ is straightforward. Going forward in time, let

$$\begin{aligned} \theta_{t|t} &\equiv E(\theta_t \,|\, Y^t, V), \\ P_{t|t-1} &\equiv Var(\theta_t \,|\, Y^{t-1}, V), \\ P_{t|t} &\equiv Var(\theta_t \,|\, Y^t, V). \end{aligned} \tag{2.23}$$

represent conditional means and variances. These are computed recursively, starting from $\bar{\theta}$ and $\bar{P}$, by iterating on

$$\begin{aligned} K_t &= (P_{t|t-1} X_t + C)(X_t' P_{t|t-1} X_t + R + X_t' C + C' X_t)^{-1}, \\ \theta_{t|t} &= \theta_{t-1|t-1} + K_t (y_t - X_t' \theta_{t-1|t-1}), \\ P_{t|t-1} &= P_{t-1|t-1} + Q, \\ P_{t|t} &= P_{t|t-1} - K_t (X_t' P_{t|t-1} + C'). \end{aligned} \tag{2.24}$$

The matrix K_t is the Kalman gain.[8] At the end of the sample, these iterations yield the conditional mean and variance for the terminal state,

$$f(\theta_T \,|\, Y^T, V) = N(\theta_{T|T}, P_{T|T}). \tag{2.25}$$

This pins down the first term in (2.22).

The remaining terms in (2.22) are derived by working backward through the sample, updating means and variances to reflect the additional information about θ_t contained in θ_{t+1}.[9] Let

$$\begin{aligned} \theta_{t|t+1} &\equiv E(\theta_t \,|\, \theta_{t+1}, Y^t, V), \\ P_{t|t+1} &\equiv Var(\theta_t \,|\, \theta_{t+1}, Y^t, V), \end{aligned} \tag{2.26}$$

[7] See Kim and Nelson (1999), ch. 8.

[8] The formula for K_t differs from that given in Anderson and Moore (1979) for the case of correlated innovations because of a difference in assumptions about the timing of innovations.

[9] Notice that the backward recursions are *not* determined by the Kalman smoother. We want the mean and variance for $f(\theta_t \,|\, \theta_{t+1}, Y^t, V) = f(\theta_t \,|\, \theta_{t+1}, Y^T, V)$. The Kalman smoother computes the mean and variance for $f(\theta_t \,|\, Y^T, V)$.

represent backward estimates of the mean and variance, respectively. Because the states are conditionally normal, these can be expressed as

$$\begin{aligned} \theta_{t|t+1} &= \theta_{t|t} + P_{t|t}P_{t+1|t}^{-1}(\theta_{t+1} - \theta_{t|t}), \\ P_{t|t+1} &= P_{t|t} - P_{t|t}P_{t+1|t}^{-1}P_{t|t}. \end{aligned} \tag{2.27}$$

Therefore the remaining elements in the (2.22) are

$$f(\theta_t \mid \theta_{t+1}, Y^T, V) = N(\theta_{t|t+1}, P_{t|t+1}). \tag{2.28}$$

Notice that the smoothed covariances depend only on the output of the Kalman filter, but the smoothed conditional means depend on realizations of θ_{t+1}. Accordingly, a random trajectory for states may be drawn from a backward recursion. First, draw θ_T from (2.25), using (2.24) to compute the mean and variance. Next, conditional on its realization, draw θ_{T-1} from (2.28), using (2.27) to compute the mean and variance. Then draw θ_{T-2} conditional on the realization of θ_{T-1}, and so on back to the beginning of the sample.

2.5.2. Gibbs Step 2: Hyperparameters Given States

Conditional on Y^T and θ^T, the innovations are observable. Under the unrestricted linear transition law, these are identically and independently distributed normal random variables, and their conditional likelihood is Gaussian. When an inverse-Wishart prior is combined with a Gaussian likelihood, the posterior is also an inverse-Wishart density,

$$p(V|Y^T, \theta^T) = IW(V_1^{-1}, T_1), \tag{2.29}$$

where

$$\begin{aligned} T_1 &= T_0 + T, \\ V_1 &= \overline{V} + \overline{V}_T, \end{aligned} \tag{2.30}$$

and $\overline{V}_T$ is proportional to the usual covariance estimator,

$$(1/T)\overline{V}_T = (1/T)\sum_{t=1}^{T}(\varepsilon_t \ v_t)(\varepsilon_t \ v_t)'. \tag{2.31}$$

The posterior degree-of-freedom parameter is the sum of the prior degrees of freedom, T_0, plus the degrees of freedom in the sample, T. The posterior scale matrix is the sum of the prior and sample sum-of-squares matrices.[10]

[10] See Gelman, Carlin, Stern, and Rubin (1995).

To sample from an inverse-Wishart distribution, we exploit two facts. First, if a matrix V is distributed as $IW(S, df)$, then V^{-1} is a Wishart matrix with scale matrix S and degrees of freedom df. Second, to simulate a draw from the Wishart distribution, we take df independent draws of a random vector η_i from a $N(0, S)$ density and form the random matrix $V^{-1} = \sum_{i=1}^{df} \eta_i \eta_i'$. Since V^{-1} is a draw from a Wishart density, V is a draw from an inverse-Wishart density.

2.5.3. Summary of the Gibbs Sampler

To summarize, the Gibbs sampler iterates on two simulations, drawing states conditional on hyperparameters and then hyperparameters conditional on states. After a transitional or "burn-in" period, the sequence of draws approximates a sample from the virtual posterior, $p_L(\theta^T, V \mid Y^T)$.

2.6. Rejection Sampling

The final step is to impose the stability condition, which is done by checking the autoregressive roots at each date and rejecting draws with roots inside the unit circle. The rejection step ensures that the posterior density puts zero probability on explosive outcomes.

To confirm the validity of this procedure, we check the conditions associated with rejection sampling.[11] The normalized target density is

$$p(\theta^T, V \mid Y^T) = \frac{I(\theta^T) p_L(\theta^T, V \mid Y^T)}{\iint I(\theta^T) p_L(\theta^T, V \mid Y^T) d\theta^T dV}. \tag{2.32}$$

To perform rejection sampling, we need a candidate density, $g(\theta^T, V)$, that satisfies three properties. The candidate density must be non-negative and well-defined for all (θ^T, V) for which $p(\theta^T, V \mid Y^T) > 0$, it must have a finite integral, and the importance ratio, $R(\theta^T, V)$, must have a known upper bound, M:

$$R(\theta^T, V) = \frac{p(\theta^T, V \mid Y^T)}{g(\theta^T, V)} \leq M < \infty. \tag{2.33}$$

A natural candidate density is the virtual posterior, $p_L(\theta^T, V \mid Y^T)$. Because this is a probability density, it is non-negative and integrates to 1. Since it is an unrestricted analog of the target density, it is also well-defined for all (θ^T, V)

[11] E.g., see Gelman, et. al. (1995), pp. 303-305.

which occur with positive probability. Finally, the importance ratio is bounded by the inverse of the probability of obtaining a stable draw from the virtual posterior,

$$\begin{aligned} R(\theta^T, V) &= \frac{I(\theta^T)}{\iint I(\theta^T) p_L(\theta^T, V \mid Y^T) d\theta^T dV} \\ &\leq \frac{1}{\iint I(\theta^T) p_L(\theta^T, V \mid Y^T) d\theta^T dV} = M. \end{aligned} \tag{2.34}$$

The denominator is the expected value of $I(\theta^T)$ under the virtual posterior, or the probability of a stable draw from the unrestricted density. M is finite as long as this probability is non-zero.

Rejection sampling proceeds in two steps: draw a trial (θ_i^T, V_i) from the virtual posterior, and then accept the draw with probability $R(\theta_i^T, V_i)/M$. Since $R(\theta_i^T, V_i)/M = I(\theta_i^T)$, the second step is equivalent to accepting the trial draw whenever it satisfies the stability condition, and rejecting it when it does not.

2.7. Beliefs About the Future

Having processed data through date T, the next step is to simulate future data and states. Conditional on hyperparameters and the current state of the system, the posterior density for future data and states is quite tractable. This density can be factored into the product of a marginal distribution for future states and a conditional distribution for future data,

$$\begin{aligned} p(Y^{T+1,T+H}, \theta^{T+1,T+H} | \theta_T, V, Y^T) = p(\theta^{T+1,T+H} | \theta_T, V, Y^T) \times \\ p(Y^{T+1,T+H} | \theta^{T+1,T+H}, \theta_T, V, Y^T). \end{aligned} \tag{2.35}$$

Because the states are Markov, the first term can be factored into

$$p(\theta^{T+1,T+H} | \theta_T, V, Y^T) = \prod_{i=1}^{H} p(\theta_{T+i} | \theta_{T+i-1}, V, Y^T). \tag{2.36}$$

Apart from the restriction on explosive autoregressive roots, θ_{T+1} is conditionally normal with mean θ_T and variance Q. Similarly, conditional on θ_{T+1}, V, and Y^T, θ_{T+2} is normally distributed with mean θ_{T+1} and variance Q, and so on. Therefore, to sample from the virtual posterior for future states, we take H random draws of v_i from the $N(0, Q)$ density and iterate on the state equation,

$$\theta_{T+i} = \theta_{T+i-1} + v_i. \tag{2.37}$$

The stability restriction is implemented in the same way as in the Gibbs sampler, by checking the autoregressive roots associated with each draw and rejecting explosive draws.

Given a trajectory for future states, all that remains is to simulate future data. The second term in (2.35) can be factored into

$$p(Y^{T+1,T+H}\,|\theta^{T+1,T+H},\theta_T,V,Y^T)=\prod_{i=1}^{H}p(y_{T+i}\,|Y^{T+1,i-1},\theta^{T+1,T+H},\theta_T,V,Y^T). \tag{2.38}$$

Conditional on θ_T,V,Y^T, and a trajectory for future states, the measurement innovation ε_{T+1} is normally distributed with mean $C'Q^{-1}v_{T+1}$ and variance $R-C'Q^{-1}C$. Hence y_{T+1} is conditionally normal with mean $(X'_{T+1}\theta_{T+1}+C'Q^{-1}v_{T+1})$ and variance $R-C'Q^{-1}C$. Similarly, ε_{T+2} is conditionally normal with mean $C'Q^{-1}v_{T+2}$ and variance $R-C'Q^{-1}C$, and so on. Therefore, to sample from (2.38), we take H random draws of ε_i from a $N(\,C'Q^{-1}v_{T+i},R-C'Q^{-1}C)$ density and iterate on the measurement equation,

$$y_{T+i}=X'_{T+i}\theta_{T+i}+\varepsilon_i,\quad i=1,...,H, \tag{2.39}$$

using lags of y_{T+i} to compute X_{T+i}.

2.8. Collecting the Pieces

Combining the results of the previous sections, (2.16) can be expressed as

$$\begin{aligned}p(Y^{T+1,T+H},\theta^{T+1,T+H},\theta^T,V\,|\,Y^T)=\;&p(\theta^T,V\,|\,Y^T)\times\\&\prod_{i=1}^{H}p(\theta_{T+i}\,|\theta_{T+i-1},V,Y^T)\times\\&\prod_{i=1}^{H}p(y_{T+i}\,|Y^{T+1,i-1},\theta^{T+1,T+H},\theta_T,V,Y^T).\end{aligned} \tag{2.40}$$

To sample from this distribution, we use the Gibbs sampler to simulate a draw from $p(\theta^T,V\,|\,Y^T)$. Then, conditional on that draw, we simulate a trajectory for future states, and conditional on both of those we simulate a trajectory for future data. This provides the raw material for our analysis.

3. Stylized Facts About the Evolving Law of Motion

We study data on inflation, unemployment, and a short-term nominal interest rate. Inflation is measured using the CPI for all urban consumers, unemployment is the civilian unemployment rate, and the nominal interest rate is the yield on 3-month Treasury bills. The inflation and unemployment data are quarterly and seasonally adjusted, and the Treasury bill data are the average of daily rates in the first month of each quarter. The sample runs from 1948.1 to 2000.4. We work with a VAR(2) specification for inflation, the logit of unemployment, and the ex post real interest rate.[12]

To calibrate the prior, we estimate a time-invariant vector autoregression using data for 1948.1-1958.4. The mean of the virtual prior, $\bar{\theta}$, is the point estimate, $\bar{P}$ is its asymptotic covariance matrix, and $\overline{R}$ is the innovation covariance matrix. To initialize the other hyperparameters, we assume that $\overline{C} = 0$ and that $\overline{Q}$ is proportional to $\bar{P}$. To begin conservatively, we start with a minor perturbation from a time-invariant representation, setting $\overline{Q} = (.01)^2 \cdot \bar{P}$. In other words, our prior is that time variation accounts for only 1 percent of the standard deviation of each parameter.[13] The prior degrees of freedom, T_0, are equal to those in the preliminary sample.

This is an informative prior, but only weakly so. Because the preliminary sample contains only 4.5 data points per VAR parameter, the prior mean is just a ballpark number and the prior variance allows for a substantial range of outcomes. As time passes, the prior becomes progressively less influential and the likelihood comes to dominate the posterior.

The simulation strategy follows the algorithm described above. Starting in 1965.4, we compute posterior densities for each year through 2000, for a total of 36 years. At each date, we perform 10,000 iterations of the Gibbs sampler, discarding the first 2000 to let the Markov chain converge to its ergodic distribution.[14] Then, conditional on those outcomes, we generate 8,000 trajectories of future data and states. Each posterior trajectory is 120 quarters long and contains information

[12] The unemployment rate is bounded between 0 and 1, and the logit transformation maps this into $(-\infty, \infty)$, which is more consonant with our Gaussian approximating model. To ensure that posterior draws for unemployment lie between 0 and 1, we simulate $logit(u_t)$ and use the inverse logit transformation. The non-negativity bound on nominal interest rates is implemented by rejection sampling.

[13] The Gibbs sampler quickly adds more time variation to the system.

[14] Recursive mean graphs suggest rough convergence, though some wiggling persists beyond the burn-in period. We checked our results by performing a much longer simulation based on data through 2000.Q4. The longer simulation involved 106,000 draws from the Gibbs sampler, the first 18,000 of which were discarded to allow for convergence. Smoothed estimates based on this simulation were qualitatively similar to the filtered estimates reported in the text. Indeed, we also performed calculations based on a burn-in period of 98,000 and found that the results were much the same.

about both short- and long-run features of the data.

3.1. Objects of Interest

We initially focus on three features of the data, long-horizon forecasts of inflation and unemployment, the spectrum for inflation, and selected parameters of a version of the Taylor rule for monetary policy. The long-horizon forecasts approximate core inflation and the natural rate of unemployment, the spectrum encodes information about the variance, persistence, and predictability of inflation, and the Taylor-rule parameters summarize the changes in monetary policy that underlie the changing nature of inflation.

We are interested in these features because they play a role in theories about the rise and fall of U.S. inflation. For example, Parkin (1993) and Ireland (1999) point out that the magnitude of inflationary bias in the Kydland-Prescott (1977) and Barro-Gordon (1983) model depends positively on the natural rate of unemployment. Taylor (1997, 1998) and Sargent (1999) argue that core inflation depends on the monetary authority's beliefs about the natural rate hypothesis, which in turn depend on the degree of inflation persistence. In particular, the model presented in Sargent (1999) imposes a definite restriction on the joint evolution of core inflation and the degree of persistence, which we discuss below. Changes in beliefs about the natural rate hypothesis should also be reflected in Taylor-rule parameters.

3.2. Core Inflation and the Natural Rate of Unemployment

Beveridge and Nelson (1981) define a stochastic trend in terms of long-horizon forecasts. For a driftless random variable like inflation or unemployment, the Beveridge-Nelson trend is defined as the value to which the series is expected to converge once the transients die out,

$$\tau_t = \lim_{h \to \infty} E_t x_{t+h}. \tag{3.1}$$

Assuming that expectations of inflation and unemployment converge to the core and natural rate as the forecast horizon lengthens, the latter can be approximated using this measure.[15] Because the posterior distributions are skewed and have fat tails, we modify the Beveridge-Nelson definition by substituting the posterior

[15] Hall (1999) recommends an unconditional mean of unemployment as an estimator of the natural rate of unemployment.

median for the mean. We approximate core inflation and the natural rate of unemployment by setting $h = 120$ quarters and finding the median of the posterior predictive density,

$$\begin{aligned} \pi_{ct} &= \text{med}_t(\pi_{t+120}), \\ u_{nt} &= \text{med}_t(u_{t+120}). \end{aligned} \tag{3.2}$$

Estimates of core inflation and the natural rate are shown in figure 3.1. The circles represent inflation, and x's refer to unemployment. According to this measure, core inflation was between 1.75 and 4 percent in the late 1960s. It rose throughout the 1970s and peaked at roughly 8 percent in 1979-80. Thereafter it fell quickly, and it has fluctuated between 2.25 and 3.25 percent since the mid-1980s. Core inflation was just shy of 3 percent at the end of 2000.

The natural rate of unemployment also rose throughout the 1970s, reaching a peak of 6.6 percent in 1980. It declined gradually in the early 1980s and fluctuated between 5.5 and 6 percent from the mid-1980s to the mid-1990s. The natural rate again began to fall after 1994 and was a bit less than 5 percent at the end of 2000.

A scatterplot, shown in figure 3.2, provides a better visual image of the association between the two. The simple correlation is 0.63, which is rather remarkable given the difficulty of measuring these components. The two series rise and fall together, in accordance with Parkin and Ireland's theory.

As a reality check for the model, figures 3.3 and 3.4 report the cyclical components of inflation and unemployment, measured by subtracting the median Beveridge-Nelson trend estimates from the actual values. We include these plots to confirm that the model captures important features of the data. The first figure shows that the estimated peaks and troughs occur at the right times and are of plausible magnitude. For example, unemployment was well above the natural rate following the recessions of 1975 and 1982. Using Okun's Law as a rule of thumb, these estimates correspond to "output gaps" of roughly 6.75 and 12.5 percent, respectively. The model also correctly predicts that the high inflation of 1974-75 and 1980-81 would be partially reversed.

Figure 3.4 plots a scatter of the cyclical components and illustrates two other characteristics of the data. The first is that the components are asymmetric, with large positive deviations occurring more often than large negative values. Second, from 1967 until 1983, there were large counter-clockwise loops in inflation and unemployment, with increases in inflation leading increases in unemployment. After 1986, the loops were smaller but still mostly counter-clockwise. The direction

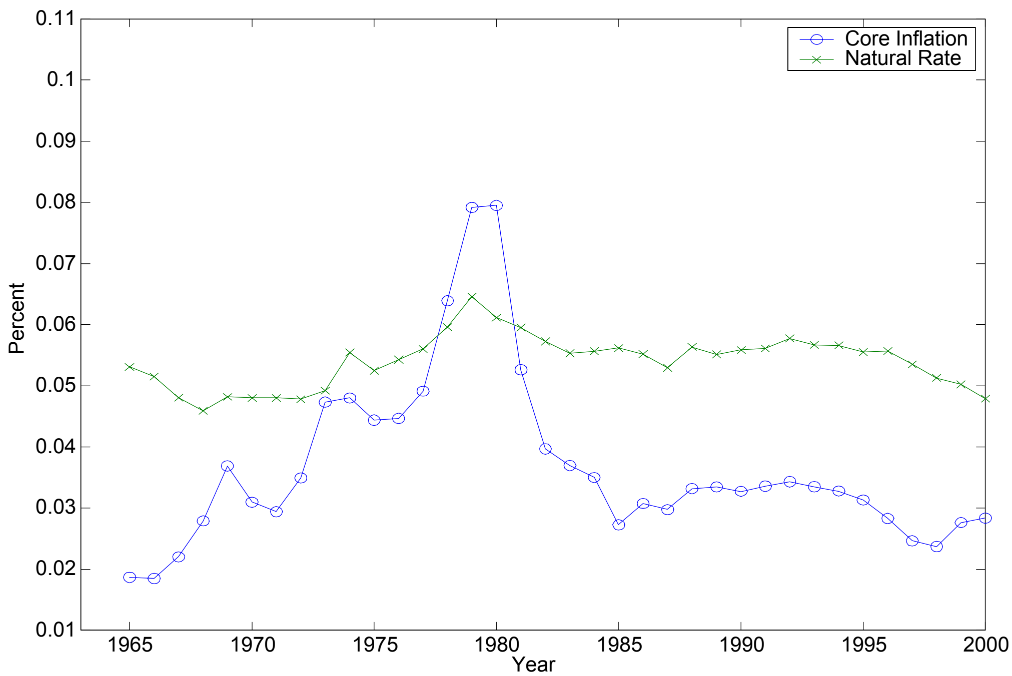

Figure 3.1: Core Inflation and the Natural Rate of Unemployment

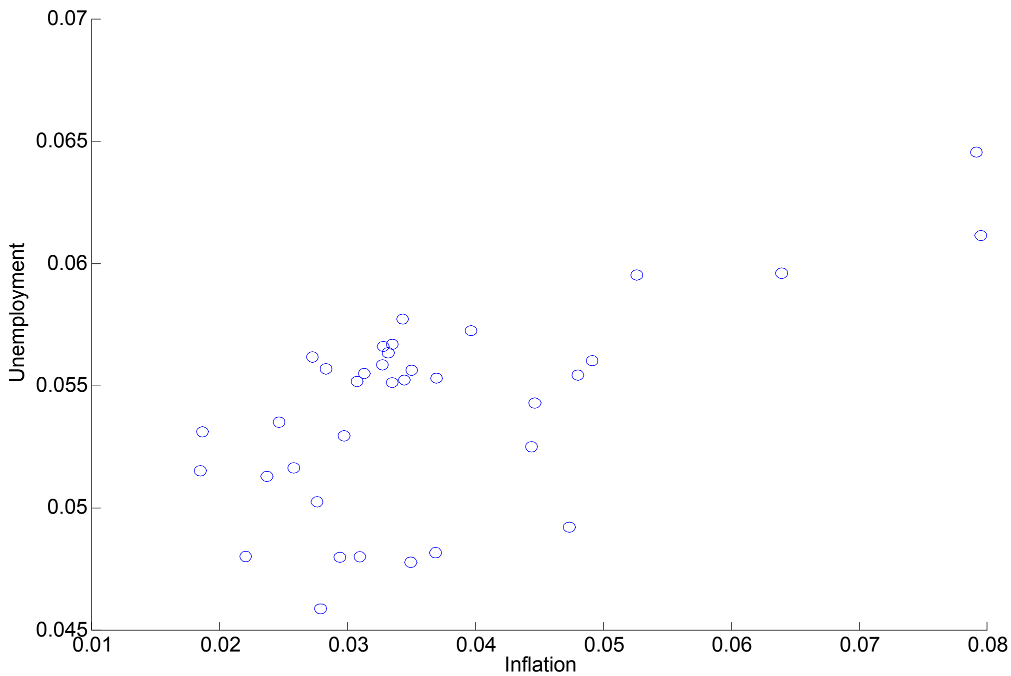

Figure 3.2: Core Inflation and the Natural Rate of Unemployment

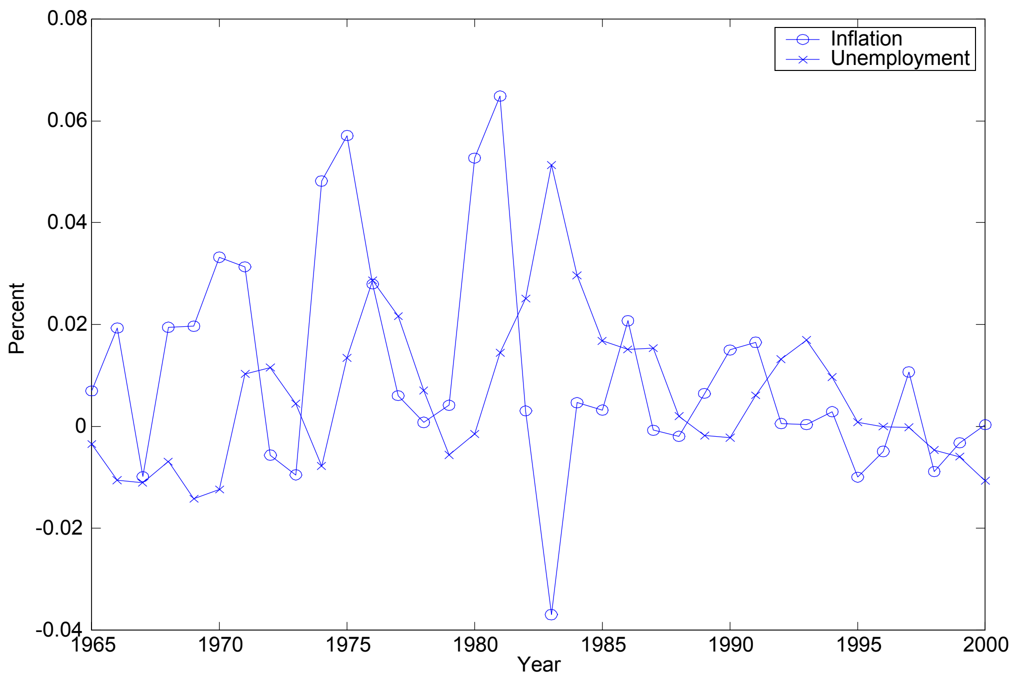

Figure 3.3: Cyclical Components of Inflation and Unemployment

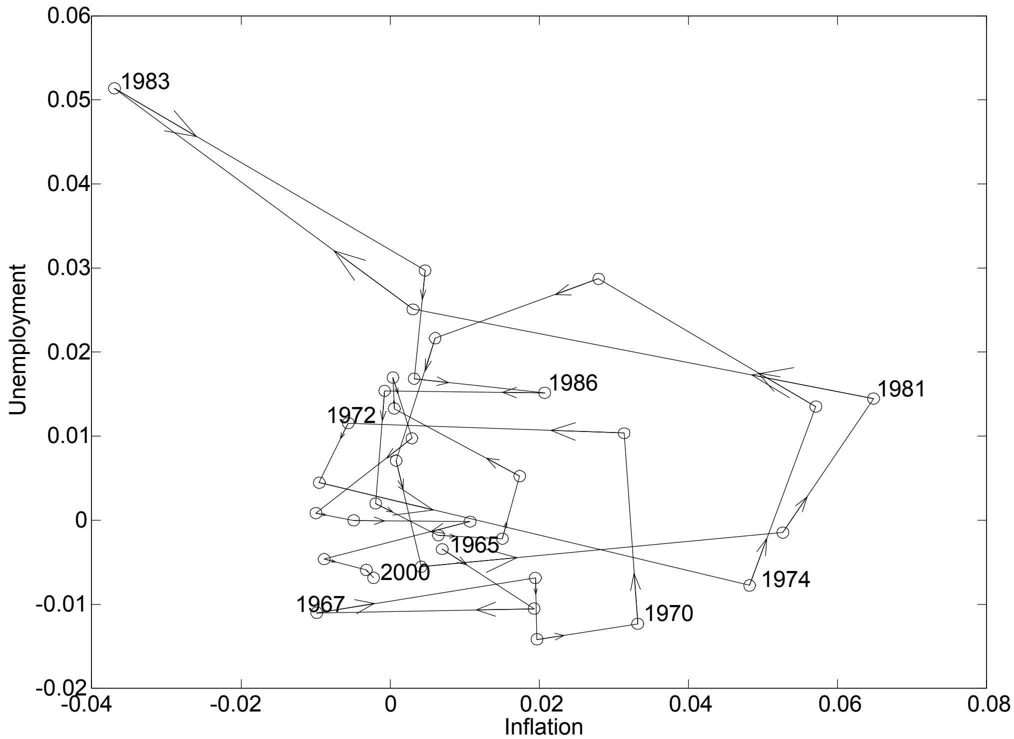

Figure 3.4: Cyclical Components of Inflation and Unemployment

of the loops is consistent with other evidence on the cyclical relation between inflation and economic activity, e.g. as summarized by Taylor (1999).

Beveridge-Nelson measures often suggest that all the variation is in the trend, a feature to which many economists object. Our model does not have this feature.

3.3. The Persistence, Variance, and Predictability of Inflation

Next we consider the evolution of the second moments of inflation. This information is encoded in the spectrum, and its evolution is illustrated in figures 3.5 through 3.7.

Figure 3.5 shows the median posterior spectrum for each year in the sample. This figure was generated as follows. For each year, we estimated a spectrum for each inflation trajectory in the posterior predictive density. Then we computed a median spectrum by taking the median of the estimates on a frequency-by-frequency basis.[16] This yields a single slice of the figure, relating power to frequency for a given year. By repeating this for each year, we produced the three-dimensional surface shown in the figure. We emphasize that these are predictive measures, which represent expected variation going forward in time. That is, the slice associated with a given year represents a prediction about how inflation is likely to vary in the future, conditional on data up to the current date.[17]

The most significant feature of this graph is the variation over time in the magnitude of low-frequency power. Since the spectral densities have Granger's (1966) typical shape, we can interpret low-frequency power as a measure of inflation persistence. According to this measure, inflation was weakly persistent in the 1960s and 1990s, when there was little low-frequency power, but strongly persistent in the late 1970s, when there was a lot. Indeed, the degree of persistence peaked in 1979-80, at the time as the peak in core inflation.

Figures 3.6 and 3.7 report results for selected years. Here, circles represent 1965, x's refer to 1979, and asterisks stand for 2000. Figure 3.6 plots the spectrum, and figure 3.7 plots its logarithm. To interpret the figures, recall that the total variance is the integral of the spectrum,

$$\sigma_\pi^2 = (1/2\pi) \int_{-\pi}^{\pi} f_{\pi\pi}(\omega) d\omega, \qquad (3.3)$$

[16] The ordinates are asymptotically independent across frequencies.

[17] We also calculated an alternative local linear approximation using the VAR representation and the mean posterior state at each date. The results were similar to those shown in the figure.

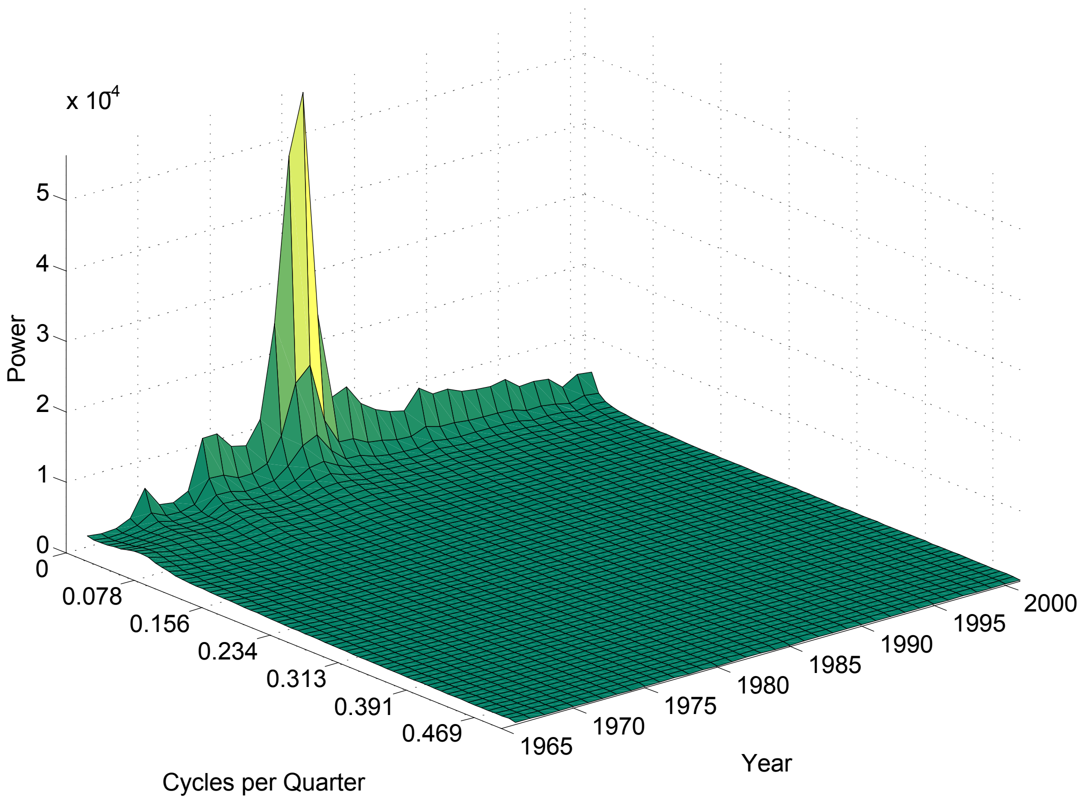

Figure 3.5: Median Posterior Spectrum for Inflation. Power is measured in basis points, the units of measurement for the variance of inflation.

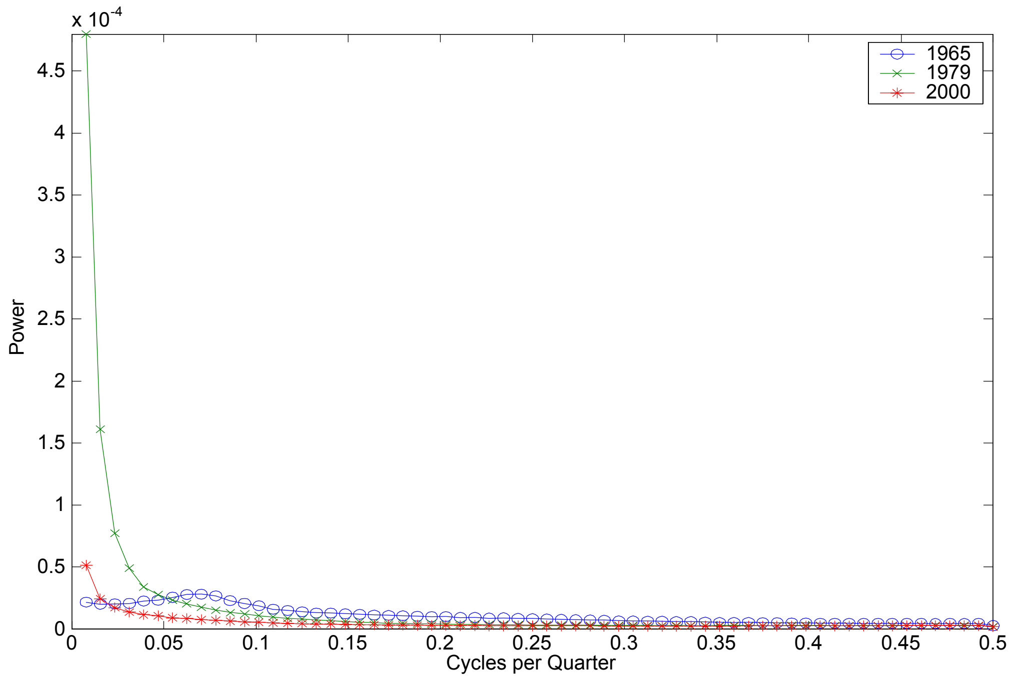

Figure 3.6: Median Posterior Spectrum for Inflation in Selected Years

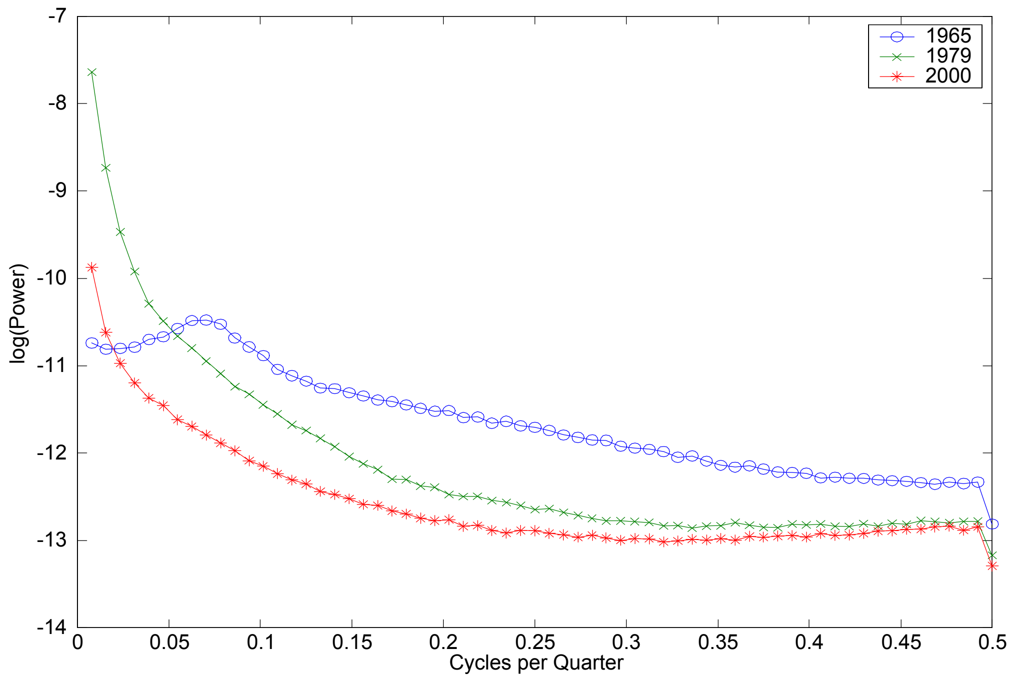

Figure 3.7: Log of the Median Posterior Spectrum for Inflation in Selected Years

and that the log of the univariate innovation variance can be expressed as the integral of the log of the spectrum,

$$\ln(\sigma_\varepsilon^2) = (1/2\pi)\int_{-\pi}^{\pi} \ln f_{\pi\pi}(\omega)d\omega. \tag{3.4}$$

The function $f_{\pi\pi}(\omega)$ is the spectrum at frequency ω, σ_π^2 is the variance, and σ_ε^2 is the error variance for one-step ahead univariate forecasts of inflation. The former measures long-run uncertainty about inflation, and the latter is a measure of short-run uncertainty.

Looking first at figure 3.6, we can say something about how the total variance has changed over time. Between 1965 and 1979, inflation became smoother but more persistent. That is, there was less variation at high and medium frequencies, especially those associated with business cycles (say 4 to 20 quarters per cycle), but more variation at low frequencies, especially those corresponding to cycles lasting 5 years or more. The increase in low-frequency power was greater in magnitude than the decrease in high-frequency power, so the total variance was greater. Thus, the increase in variance during the late 1960s and 1970s reflected an increase in inflation persistence.

Between 1979 and 2000, the spectrum for inflation fell at all frequencies, and therefore so did the total variance. But the decline in power was greatest at low frequencies, especially at those greater than 20 quarters per cycle. In other words, the diminished degree of inflation persistence accounted for most of the decline in variance in this period. Thus the evolution of the variance has been closely associated with that of inflation persistence. Inflation became more persistent and more variable in the 1970s and less persistent and less variable in the 1980s and 1990s.

Figure 3.7 is relevant for short-term forecasting and tells a somewhat different story. The increase in the log of low-frequency power between 1965 and 1979 was smaller in magnitude than the decrease in the log of high-frequency power. Thus, although inflation became more persistent and more variable during the 1970s, it also became easier to predict one quarter ahead. In other words, although there was more long-term uncertainty in 1979, there was actually less short-term uncertainty. Between 1979 and 2000, the log-spectrum fell at all frequencies, and inflation became even easier to forecast one-quarter ahead. By 2000 there was less uncertainty at both long and short horizons.

The next two figures provide more information about prediction errors. Figure 3.8 is a multivariate analog of figure 3.7 and is related to the total prediction

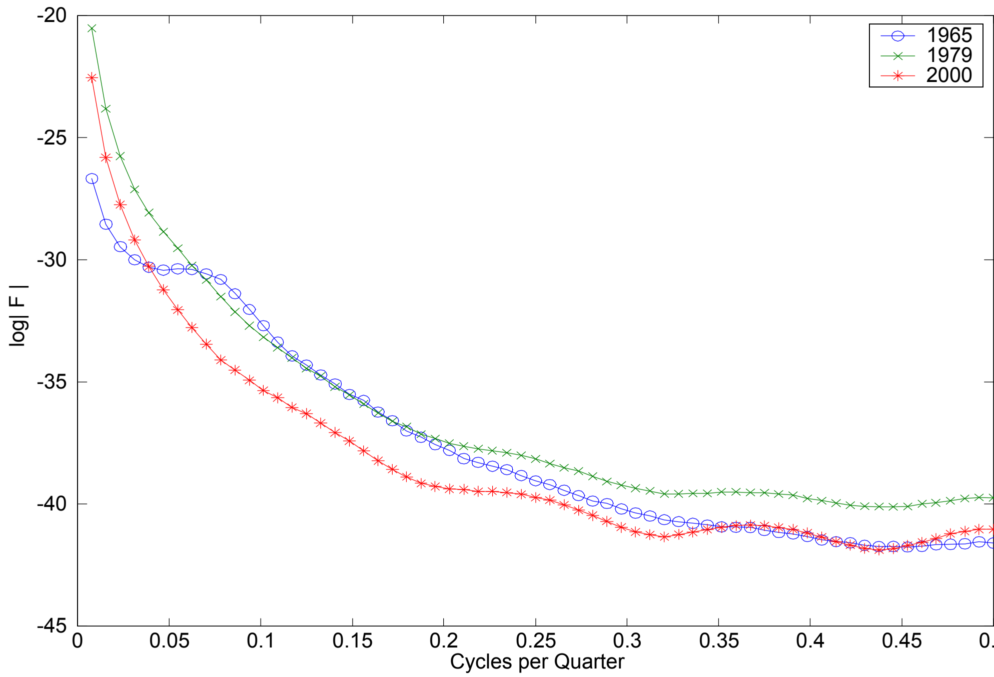

Figure 3.8: Log-Determinant of the Mean Posterior Spectral Density Matrix in Selected Years

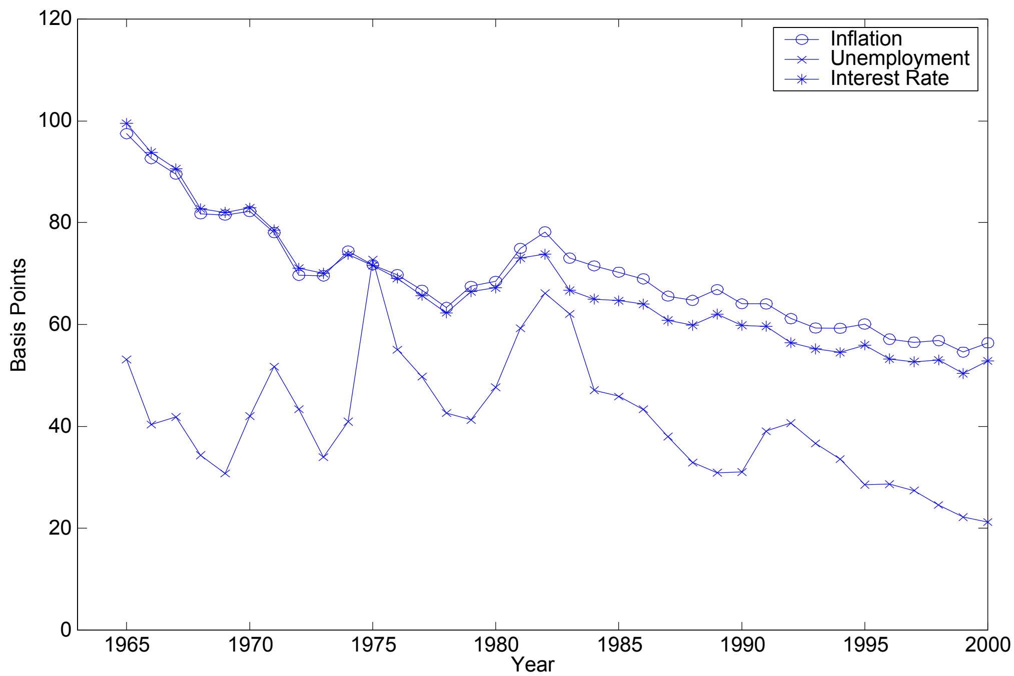

Figure 3.9: Standard Deviation of One-Step Ahead VAR Prediction Errors

variance for the system. To interpret this figure, recall that the total prediction variance, $|V_{\varepsilon\varepsilon}|$, for a vector time series y_t can be expressed in terms of the log of the determinant of the spectral density,

$$\ln(|V_{\varepsilon\varepsilon}|) = (1/2\pi)\int_{-\pi}^{\pi} \ln|F_{yy}(\omega)|d\omega, \tag{3.5}$$

where $V_{\varepsilon\varepsilon}$ is the covariance matrix for innovations based on the history of y_t and $F_{yy}(\omega)$ is the spectral density matrix. Whittle (1953) interprets $|V_{\varepsilon\varepsilon}|$ as a measure of the total random variation entering the system at each date.

Unlike the univariate measure, the total prediction variance increased between 1965 and 1979. For the system as a whole, there was only a slight decrease in variation at business cycle frequencies, and this was more than offset by a substantial increase in variation at low and high frequencies. Between 1979 and 2000, the system became more predictable, with $\ln|F_{yy}(\omega)|$ falling at all frequencies. This more than reversed the increase in the earlier period. By the end of 2000, the total prediction variance was 40 percent smaller than in 1979 and 30 percent smaller than in 1965. Thus, for the system as a whole, the degree of short-term uncertainty has fallen substantially.

Figure 3.9 reports the variance of VAR forecast errors over the period 1965-2000 and provides more detail about the evolution of short-run uncertainty. At each date, the posterior prediction error variance was computed by averaging across realizations of the posterior predictive density, one-quarter ahead. For inflation and ex post real interest rates, there has been a downward trend in short-term uncertainty since 1965, punctuated by an increase in 1974 and again in the 1978-82. According to this measure, the VAR innovation variance for inflation fell by 21 percent between 1979 and 2000 and by 42 percent for the period as a whole. In contrast, the forecast error variance for unemployment fluctuated until the early 1980s, rising and falling with the business cycle. Since then it has fallen steadily to less than one-third its peak level. Changes in short-run uncertainty about unemployment account for much of the rise and fall of the total prediction variance.[18]

Finally, in figures 3.10 and 3.11, we relate changes in core inflation to the evolution of the variance and degree of persistence of inflation. Figure 3.10 plots

[18] Although our model assumes that V is constant, the figures illustrate that filtered estimates do shift little by little over time, thus introducing a limited degree of variation in shock variances. This variation may reflect a transient adaptation to the kind of shifts emphasized by our discussants.

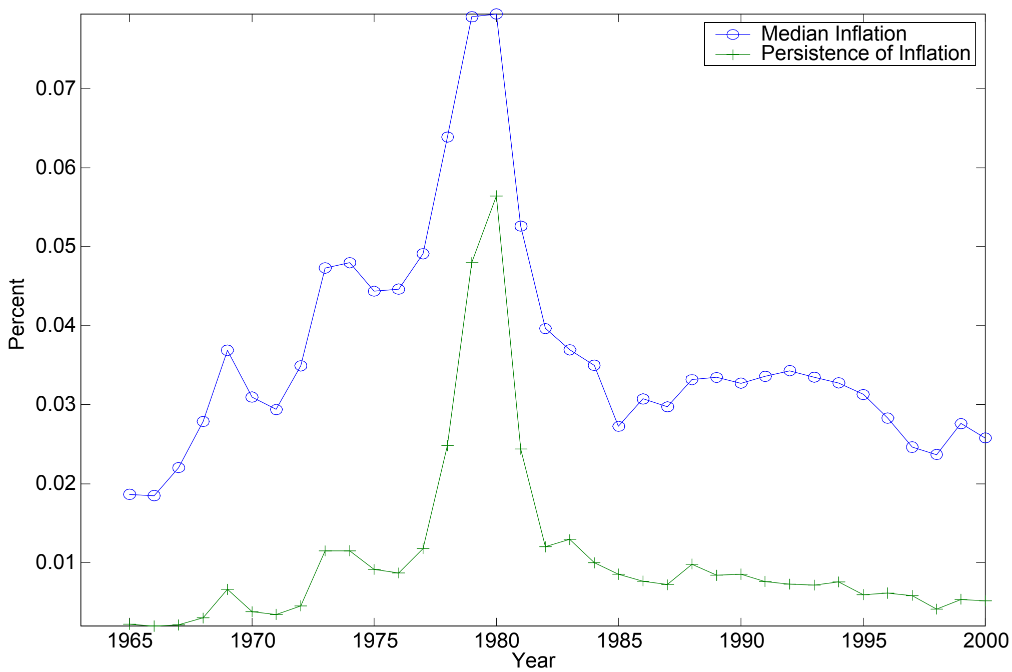

Figure 3.10: Core Inflation and Inflation Persistence

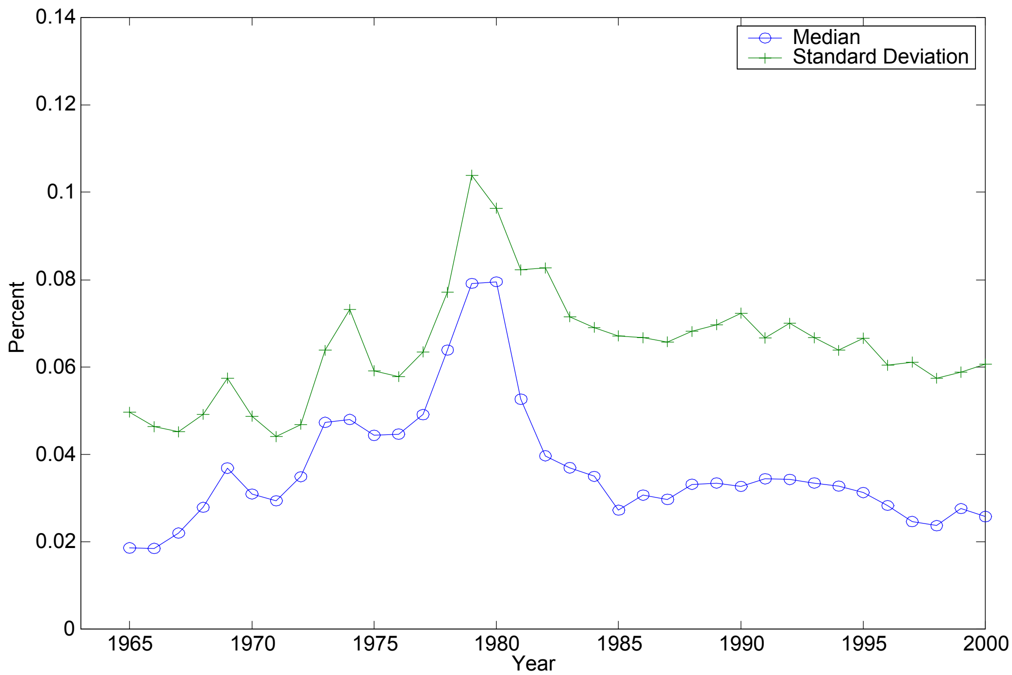

Figure 3.11: Core Inflation and the Standard Deviation of Inflation, 30 Years Ahead

core inflation and the spectrum at frequency zero, which summarizes the degree of persistence. The two are very closely related. Both rose in the 1960s and 1970s and both fell during and after the Volcker disinflation. The simple correlation is 0.915.

Because persistence contributes to variance, core inflation also covaries positively with the long-horizon standard deviation of inflation, as shown in figure 3.11.[19] Again, both measures rose during the 1970s and fell during the 1980s and 1990s. The correlation between the mean and standard deviation is 0.783. This is a bit lower than the previous correlation because the variance includes changes in both low- and high-frequency power, and the latter are less highly correlated with changes in core inflation. Thus the well-known positive correlation between the mean and variance of inflation reflects an even stronger correlation between the mean and degree of persistence.

3.4. Taylor Rule Parameters

At the end of the day, we hope to interpret the evolution of inflation dynamics in terms of the changing behavior of central bankers. Accordingly, we also investigate the evolution of the parameters of a Taylor rule.

A simple form of the Taylor rule posits that the central bank's nominal interest target, i_t^*, varies positively with inflation and inversely with unemployment,

$$i_t^* = (r^* + \pi^*) + \beta(\pi_{t-1} - \pi^*) + \gamma(u_{t-1} - u^*), \tag{3.6}$$

where π^*, u^* and r^* represent target values for inflation, unemployment, and the real interest rate, respectively. The lags in the relationship reflect the fact that current observations on inflation and unemployment are often unavailable to policy makers, especially early in the quarter.[20] Therefore decisions are based on lagged values of inflation and unemployment. The basic Taylor rule is usually augmented with a policy shock, η_t, and a partial adjustment formula to allow for interest rate smoothing,

$$\Delta i_t = \rho(L)(i_t^* - i_{t-1}) + \eta_t. \tag{3.7}$$

Cast in this form, the Taylor rule can be represented as the interest rate equation in a vector autoregression for inflation, unemployment, and nominal interest rates.

[19] We focus on the long-horizon variance, $\text{var}_t(\pi_{t+120})$, in order to let the transients die out.

[20] This is relevant in our case because the interest rate is sampled in the first month of the quarter.

In an alternate form of the Taylor rule, decisions about the ex ante real interest rate depend on lags of inflation, unemployment, and ex post real rates,

$$i_t - E_{t-1}\pi_t = \mu + \beta(L)\pi_{t-1} + \gamma(L)u_{t-1} + \rho(L)(i_{t-1} - \pi_{t-1}) + \eta_t. \tag{3.8}$$

By substituting $\pi_t = E_{t-1}\pi_t + \varepsilon_{\pi t}$, this form can be cast as the real interest equation in a vector autoregression for inflation, unemployment, and ex post real rates, with a composite innovation consisting of policy shocks and inflation prediction errors,

$$i_t - \pi_t = \mu + \beta(L)\pi_{t-1} + \gamma(L)u_{t-1} + \rho(L)(i_{t-1} - \pi_{t-1}) + (\eta_t - \varepsilon_{\pi t}). \tag{3.9}$$

This is the form of the Taylor rule that we shall study.[21] In response to our discussants, we concede that it is controversial to interpret the systematic part of the monetary policy rule as the projection of real interest rates only on *past* information. By orthogonalizing an innovation covariance matrix in a particular order, many studies attribute part of the contemporaneous covariance among innovations to the monetary rule (i.e., the rule for setting interest rates responds to contemporary information). We also recognize that the shapes of impulse response functions of macroeconomic aggregates to the monetary policy shock can depend sensitively on how much of the contemporaneous innovation volatility is swept into the monetary shock. In defense of our choice, we note that among others McCallum and Nelson (1999) doubt that monetary authorities have timely and reliable enough reports to let them respond to what the vector autoregression measures as contemporaneous information.[22]

The literature on monetary policy rules emphasizes several aspects of central bank behavior. We focus on two elements that are especially relevant to the evolution of the law of motion for inflation. One concerns the evolution of target inflation, π^*, and the other concerns the evolution of the degree of activism.

The value of target inflation cannot be identified from the interest rate equation alone. But assuming that the central bank adjusts interest rates so that inflation eventually converges to its target, this can be estimated by computing long-horizon forecasts using the entire vector autoregression. Under this assumption, target and core inflation are synonymous. Evidence on this feature of the policy rule is reported above, in figure 3.1.

[21] Actually, we substitute the logit of unemployment for unemployment.

[22] It would have been possible for us to condition on contemporaneous information by using the time t estimate of the R component of V to orthogonalize R as desired, though we have not done so in this paper.

Another important issue concerns whether a rule is "activist" or "passivist," a distinction that bears on the determinacy of equilibrium (e.g., see Clarida, Gali, and Gertler 2000). A policy rule is activist if, other things equal, the central bank increases the nominal interest rate more than one-for-one in response to an increase in inflation, so that the real interest rate increases. A passivist central bank adjusts the nominal interest rate by one-for-one or less, so that the real interest remains constant or falls as inflation rises. In the real interest version of the Taylor rule, the degree of activism can be measured by

$$A = \frac{\beta(1)}{1 - \rho(1)}. \tag{3.10}$$

A policy rule is activist if $A > 0$.

Because our version of the Taylor rule is the real interest equation in the vector autoregression, the posterior density for the activism coefficient can be computed directly from the posterior density for the states. The output of the Gibbs sampler at date t includes the terminal state, θ_t, and for each draw of the terminal state we calculate the implied value for A. Conditional on data up to date t, this measures the degree of activism that would be forecast going forward from date t.

Posterior beliefs about A are illustrated in figure 3.12. Because of outliers in the posterior density, the figure graphs the posterior median and interquartile range.[23] The figure has two salient features. First, as reported by Clarida, et. al., there have been important changes in the degree of activism over time. Judging by the posterior median, which is marked by circles, the degree of activism declined in the late 1960s and was approximately neutral in the early 1970s. For the remainder of the 1970s, the rule was decidedly passive, allowing real interest rates to fall as inflation rose. Monetary policy started becoming activist in 1981 and continued to grow more activist until the end of Volcker's term. During the first half of Greenspan's term, policy drifted toward a less active stance, perhaps reflecting the "opportunistic" approach to disinflation. But policy has again grown more activist since 1993, surpassing the peak achieved at the end of the Volcker years.

The second notable feature concerns the dispersion of beliefs about the degree of activism. Judging by the interquartile range, beliefs were tightly concentrated only in the 1970s, when monetary policy was passive. At that time, there seemed to be little doubt, for better or worse, about how the Fed was doing business. The periods before and after both involve more uncertainty about the degree of activism. In the 1960s, the lower end of the interquartile range straddled the

[23] The outliers result from division by $1 - \rho(1)$, which sometimes takes on values close to zero.

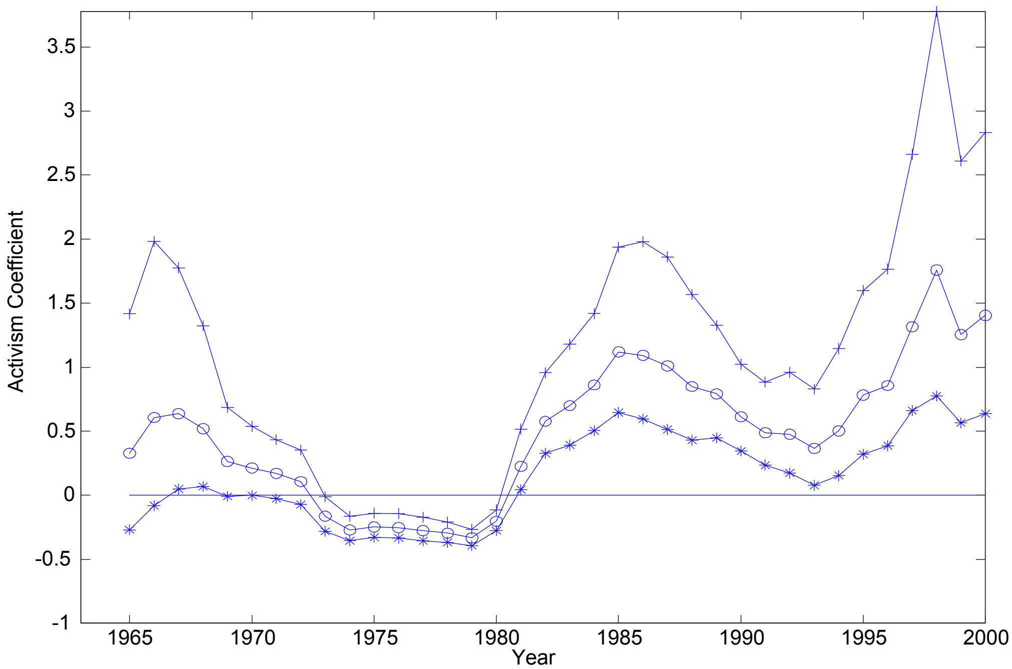

Figure 3.12: Posterior Median and Interquartile Range for the Activism Coefficient

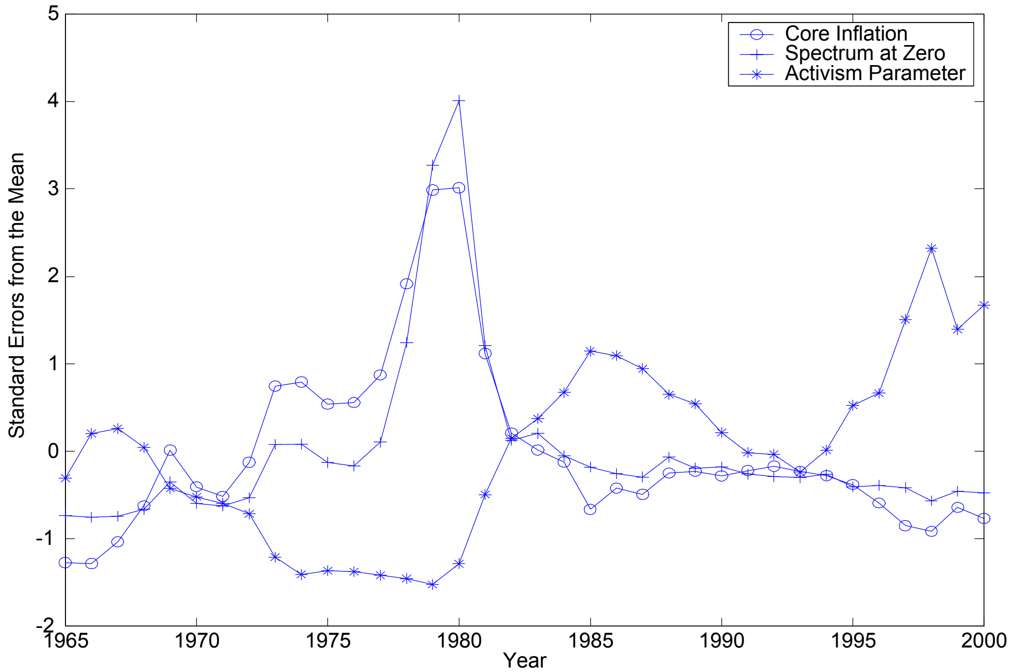

Figure 3.13: Core Inflation, Inflation Persistence, and Policy Activism

boundary of the activist region. In the Volcker-Greenspan years, the interquartile range was wider but safely within the activist region.

Figure 3.13 shows how the activism parameter has covaried with core inflation and the degree of inflation persistence.[24] The latter both increased during the 1970s experiment with a passivist monetary rule, and they both fell in the 1980s and 1990s as policy became more activist. The correlation between the degree of activism and core inflation is -0.69 over the full sample and -0.87 in the Volcker-Greenspan era. Similarly, the correlation between the activism and persistence measures is -0.46 over the full sample and -0.76 in the Volcker-Greenspan years. Thus, as one might expect, there is an inverse relation between the degree of activism on the one hand and core inflation and inflation persistence on the other.

4. Testing the Natural Rate Hypothesis

Figures 4.1 through 4.3 summarize the consequences of implementing econometric tests of the natural rate hypothesis along the lines of Solow (1968), Tobin (1968), Gordon (1970), and many others. They tested the natural rate hypothesis by regressing inflation on its own lags along with current and lagged values of unemployment,

$$\pi_t = \beta_0 + \beta_1(L)\pi_{t-1} + \beta_2(L)u_t + \varepsilon_t. \tag{4.1}$$

They interpreted the condition $\beta_1(1) = 1$ as evidence in favor of the natural rate hypothesis and $\beta_1(1) < 1$ as evidence in favor of a long-run trade-off.[25]

The outcomes of recursive natural rate tests are shown in figure 4.1. The initial estimates are based on data from 1948 through 1964, allowing for lags at the beginning of the sample. On the right-hand side of equation (4.1), we include two lags of inflation along with the current value and two lags of unemployment.

[24] The variables are measured in standard units in order to put them on a common basis.

[25] The thought experiment in play imagines the consequences of a permanent increase in expected inflation, which is proxied by the lagged inflation terms on the right-hand side. In order for this to be neutral in the long run, it must be the case that this has a one-for-one effect on actual inflation, so that $\beta_1(1) = 1$. Assuming that current unemployment is predetermined with respect to current inflation, this regression can be estimated by least squares. King and Watson (1997) point out that the last assumption follows from the structure of vintage 1960s Keynesian models, in which unemployment and inflation were determined in a block recursive fashion. Unemployment was determined by aggregate demand and Okun's law. Taking unemployment as given, inflation was determined by a Phillips curve relation for wages and a mark-up equation for prices.

Starting in 1965:1, new data are added one quarter at a time, and $\hat{\beta}_1(1)$ and its t-ratio are updated using the Kalman filter. The figure plots the resulting sequence of t-statistics for $\hat{\beta}_1(1) - 1$. Points marked with a circle represent OLS estimates, and those marked with a diamond represent discounted least squares (DLS) estimates. For the latter, the gain parameter was $g_t = \max(1/t, 1/120)$.[26] The horizontal line marks the one percent critical value for a one-sided test.

Sargent (1971) pointed out that this approach is valid only if the sample used to estimate $\beta_1(1)$ contains permanent shifts in inflation. Otherwise the data are uninformative for the thought experiment, and $\beta_1(1)$ could be less than 1 even if there were no long-run trade-off. Thus, as the degree of inflation persistence in the sample varies over time, so too will outcomes of the test.

Early versions of the test, based on samples in which there was little inflation persistence, found estimates of $\beta_1(1) < 1$ and were interpreted as evidence in favor of a long-run trade-off. As shown in the figure, the natural rate hypothesis was strongly rejected through 1973. Later versions were based on samples containing more inflation persistence, and they fail to reject long run neutrality. Indeed, from the mid-1970s until the mid-1980s there was very little evidence against long run neutrality. Since then, as the degree of inflation persistence has fallen, evidence against the natural rate hypothesis has grown.

Figures 4.2 illustrates the relation between inflation persistence and outcomes of the test.[27] The figure confirms that the test statistic is positively related to the degree of persistence, though the relation is nonlinear. Once there was enough persistence to identify the long run trade-off parameter, the test began to accept long run neutrality, and further increases in persistence no longer increased the t-ratio. Figure 4.3 shows that the test statistic is also positively related with core inflation. Without alterations, the model of Sims (1988), Chung (1990), and Sargent (1999) cannot explain that pattern. In that model, persistence rises and the natural rate hypothesis is learned as inflation falls, so the model predicts an inverse relation between core inflation and the outcome of the test. The pattern shown in figure 4.3 is more consistent with an alternative story, in which the upward drift in inflation taught the government to accept the natural rate hypothesis via the Solow-Tobin test.

[26] There are only minor differences between the two estimators within the sample, because until recently $1/t > 1/120$. The distinction between constant and decreasing gain estimators matters more when we consider the likely outcomes of future tests.

[27] These figures refer to discount least squares estimates, but the results for OLS estimates are essentially the same.

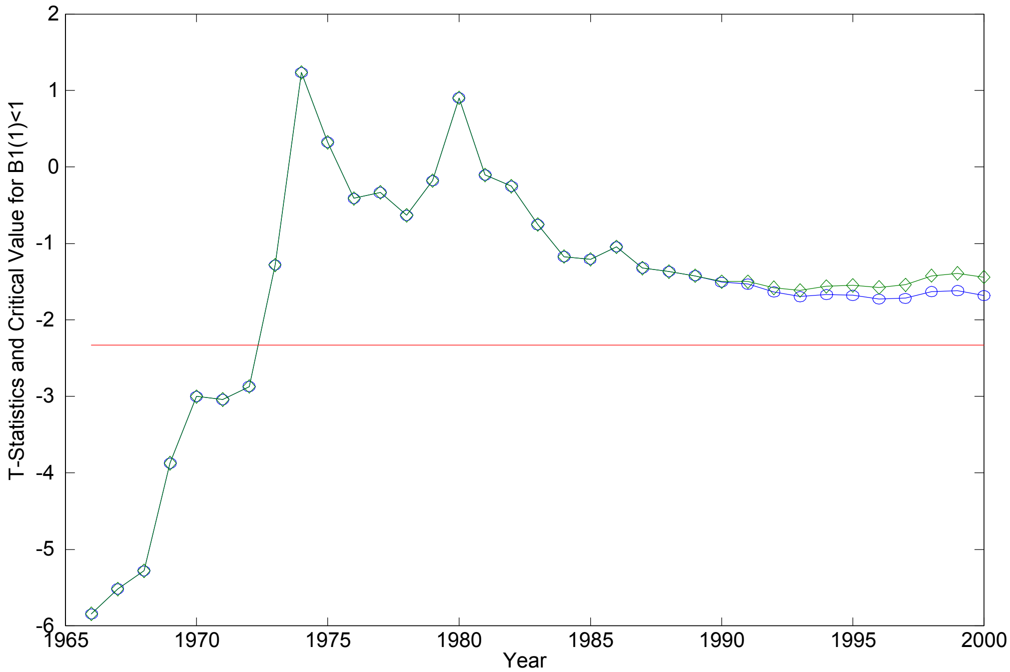

Figure 4.1: Recursive Tests of the Natural Rate Hypothesis. Circles represent recursive least square estimates with decreasing gain, and diamonds illustrate constant-gain estimates.

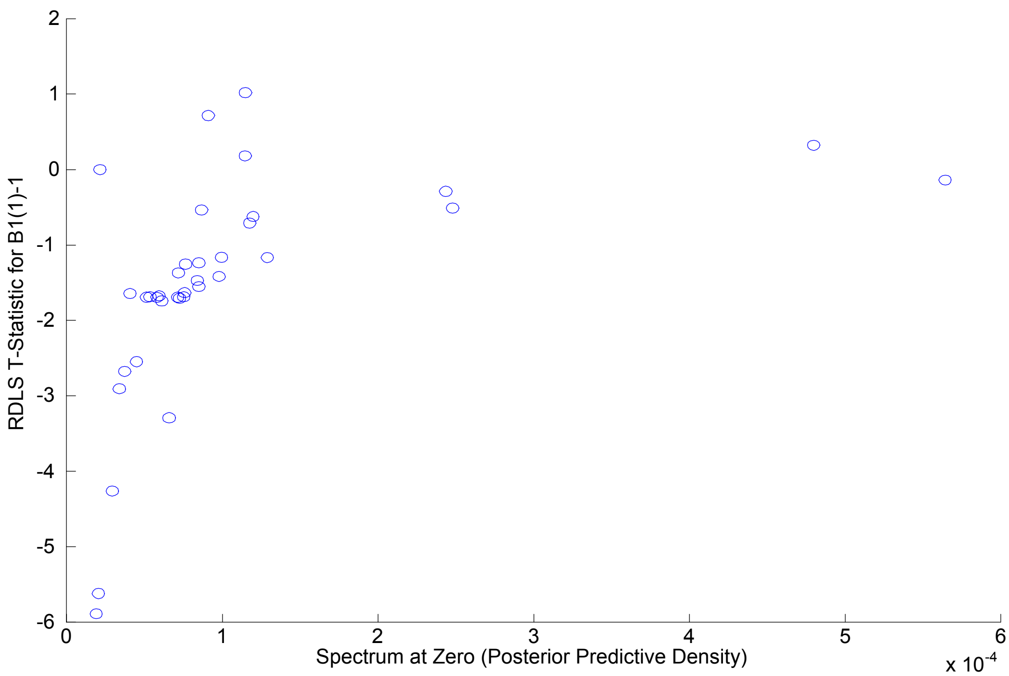

Figure 4.2: Inflation Persistence and NRH Test Statistics

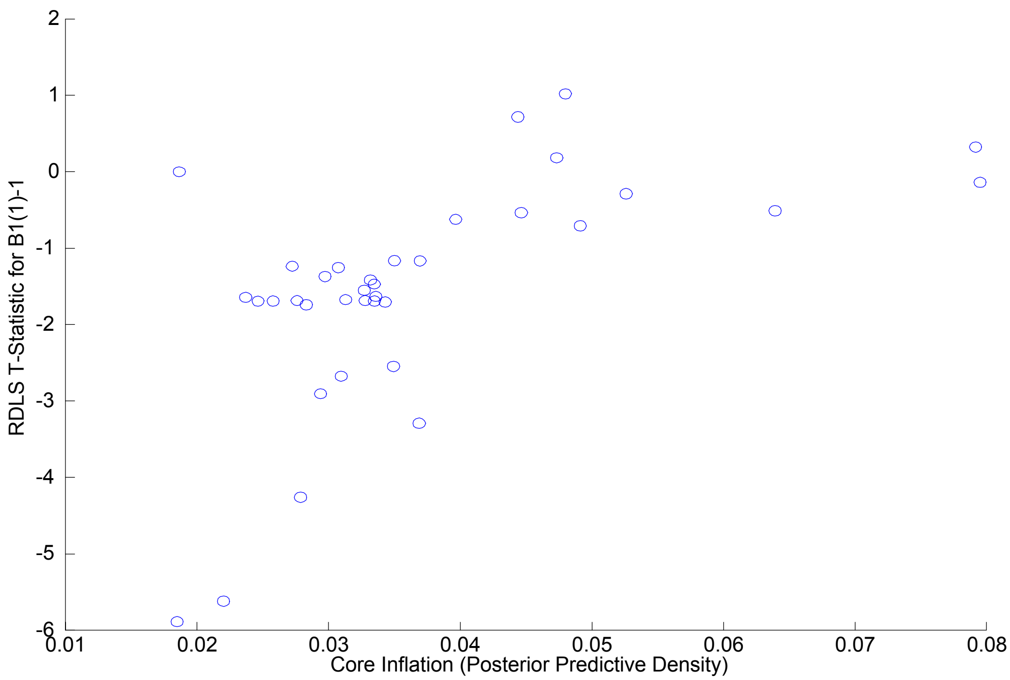

Figure 4.3: Core Inflation and NRH Test Statistics

Thus, though the Solow-Tobin procedure provided a valid test of the natural rate hypothesis only when inflation had become sufficiently persistent, by the mid 1970's inflation had become persistent enough to let the test detect the natural rate. Therefore the Solow-Tobin econometric procedures gave policy makers information that should have caused them to stabilize inflation if they had the preferences attributed to them, for example, by Kydland and Prescott (1977). For when a policy maker solves the problem of minimizing an expected discounted sum of a quadratic loss function in inflation and unemployment subject to a Phillips curve like (4.1), and when the policy maker accepts the natural rate hypothesis in the form in which Solow and Tobin cast it, then for discount factors large enough, the policy maker will soon push average inflation to zero.[28] When Volcker took control, the advice quickly to push inflation toward zero came even from those models and optimal control exercises that inadequately accounted for the Lucas critique, because they rested on the Solow-Tobin test.

However, the strong inflation persistence that induced the Solow-Tobin test to detect the natural rate in the mid 1970s depended on the monetary authority's having recently allowed inflation to drift upward, perhaps in response to its earlier erroneous views about an exploitable tradeoff. If the government's success in lowering inflation created lower persistence in inflation, the Solow-Tobin test could one day again point to an exploitable tradeoff that would tempt later monetary authorities to use inflation to fight unemployment. That possibility has worried John Taylor and others, an issue to which we now turn.

5. Taylor's Warning About Recidivism

Recently, John Taylor (1998) has warned about recidivism on the natural rate hypothesis. Taylor notes that inflation is lower and more stable in the current monetary regime, and he points out that as such data accumulate, erroneous econometric tests of long-run neutrality may again begin to suggest the existence of a trade-off. To the extent that the tests undermine confidence in the natural rate hypothesis, they could also undermine support for a low inflation policy. In this section, we offer quantitative evidence to back up Taylor's warning. The evidence is based on the posterior predictive density conditioned on data through the end of 2000. We use this to make predictions about the probability of rejecting

[28] This is a version of the control problem described by Phelps (1967) and Sargent (1999). Long ago, Albert Ando pointed out that good macroeconometric models had confirmed the absence of a long-run inflation-unemployment tradeoff by the early or mid 70's.

the natural rate hypothesis going forward in time.

Figure 4.1 suggests that Taylor's concern has some merit, because by the end of the sample conventional tests were close to rejecting $\beta_1(1) = 1$ against $\beta_1(1) < 1$ at the 5 percent level. The in-sample evidence is marginal,[29] however, and it is an open question whether stronger evidence will emerge as data from a low-inflation regime accumulate. To address this question, we compute the posterior predictive density of natural rate t-ratios going forward in time from 2000.4. Then we calculate the probability, conditioned on what we know now, of rejecting the natural rate hypothesis at various dates in the future. In this way, we can quantify the risk of backsliding.

Let $\hat{\tau}^{T+1,T+H}$ represent a potential future sequence of recursive t-statistics for $\beta_1(1) - 1$,

$$\hat{\tau}^{T+1,T+H} = [\hat{\tau}_{T+1}, ..., \hat{\tau}_{T+H}]'.$$

We want to make statements about how these sequences are likely to evolve. From a Bayesian perspective, the natural way to proceed is to compute the posterior predictive density for these sequences,

$$p(\hat{\tau}^{T+1,T+H}|Y^T). \tag{5.1}$$

To sample from this density, we start with the posterior predictive density for inflation and unemployment and then exploit the fact that t-statistics are deterministic functions of the data.[30] Hence we can write

$$p(\hat{\tau}^{T+1,T+H}|Y^T) = p(g(Y^{T+1,T+H}, Y^T)|Y^T), \tag{5.2}$$

where the function $g(\cdot)$ is nothing more than the output of the recursive least squares algorithm initialized with estimates through date T. To draw a realization from (5.2), we first draw a trajectory for future inflation and unemployment from their posterior predictive density and then apply the Kalman filter to compute the associated sequence of test statistics. The probability that the test will reject at some future date h is

$$\int_{-\infty}^{c(\alpha)} p(\hat{\tau}_h|Y^T) d\hat{\tau}, \tag{5.3}$$

[29]In our opinion, strong rejections will be needed to reverse the consensus in favor of the natural rate hypothesis.

[30]Remember, from a Bayesian perspective $\beta(1)$ is random and $\hat{\beta}(1)$ is deterministic.

where $c(\alpha)$ is the Normal critical value corresponding to a one-sided test of size α. In terms of our sampling strategy, this is the fraction of simulated trajectories in which $\hat{\beta}_1(1)$ is significantly less than 1 at date h, where significance is determined by the usual classical criterion. Thus, we are offering a Bayesian interpretation of judgments based on a classical procedure.

Figure 5.1 reports results for a constant-gain estimator. The results for a recursive OLS estimator are similar. We focus on the constant-gain estimator because this holds the effective sample size constant as data accumulate. Thus the increased probability of rejection does not follow simply from an increase in the number of observations.

As the figure shows, the probability of rejection remains small in the first two years of the forecast. But then it increases quickly, reaches 50 percent within 9 years, and approaches 85 percent in 20 years. The increasing probability of rejection reflects the changing nature of inflation-unemployment dynamics along with the fact that data from new and old regimes are being mixed in different proportions. As time moves forward, data from the old high-inflation, strong-persistence regime are discounted more heavily, and data from the new low-inflation, weak-persistence regime increasingly dominate the sample. The identifying information from the 1970s is lost little by little, and the properties of the Volcker-Greenspan era come more and more into play. This confirms an element of Taylor's warning, that the Solow-Tobin test may once again begin to suggest the existence of a trade-off.

6. Concluding Remarks

This paper has used a vector autoregression with random coefficients to measure parameter drift in U.S. inflation-unemployment-interest rate dynamics. We construct our model to focus on parameter drift because we are sympathetic to the theoretical views expressed in Lucas (1976) and Sargent (1999), which lead us to suspect that evolution in the monetary policy authority's view of the world will make the systematic part of a vector autoregression drift. We have taken seriously our model's description of four sources of uncertainty about the future,[31] and have used computer intensive Bayesian methods to take those uncertainties into account. We use the model to develop a number of stylized facts about the

[31] These are: (1) the unknown current location of the VAR coefficients, (2) the unknown covariance matrix of innovations to VAR coefficients and equations, (3) the future evolution of the VAR coefficients, and (4) the stream of future shocks to the VAR equations.

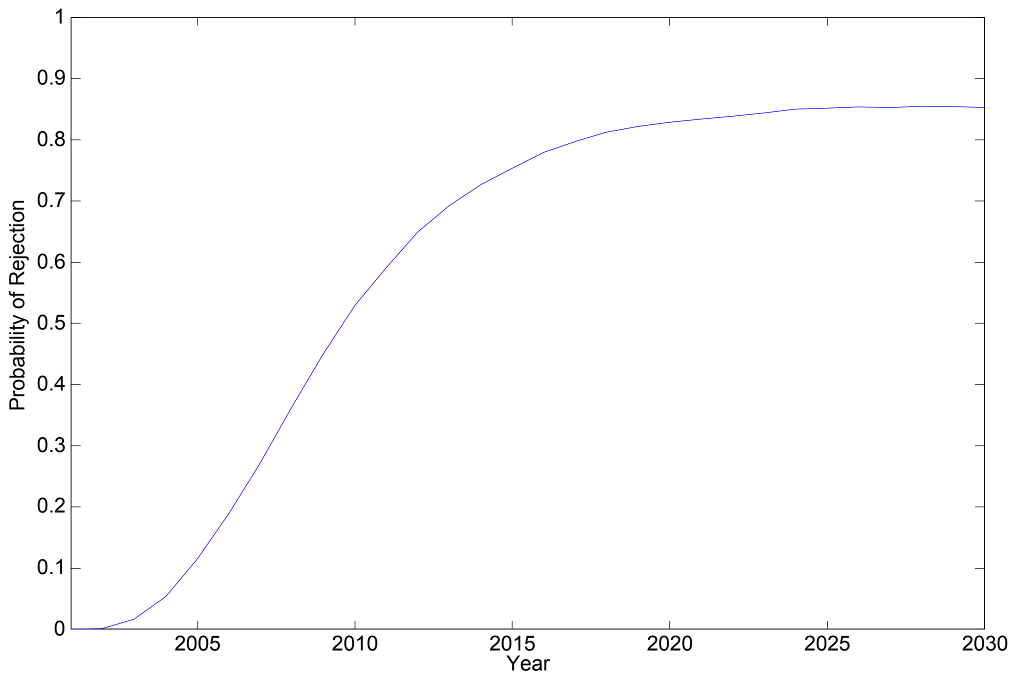

Figure 5.1: Probability of Rejecting the Natural Rate Hypothesis, Calculated from the Posterior Predictive Distribution for the Constant-Gain Estimator.

evolution of post-war U.S. inflation and relate them to important issues about learning to detect the natural rate hypothesis using imperfect tests, and how the evolving results from those tests were associated with evolution in a description of a monetary policy rule (a 'Taylor rule'). Among other things, we find that the mean and persistence of inflation are strongly positively correlated; that the persistence of inflation is positively associated with statistics that have been used to test for accepting the natural rate hypothesis; that evolving measures of policy activism in fighting inflation broadly point to more activism with a lag somewhat after test statistics began accepting the natural rate hypothesis; and that recently the degree of persistence in inflation has been drifting downward as inflation has come under control.

We also study John Taylor's warning about recidivism toward an exploitable trade-off between inflation and unemployment. Unfortunately, our statistical model confirms Taylor's concerns. Our model predicts that as observations of lower, more stable inflation accumulate, econometric evidence against the natural rate hypothesis is likely to develop.[32] Against this evidence, we hope that policy makers do not succumb again to the temptation to exploit the Phillips curve.

7. Appendix: A Nonlinear Transition Equation

Our numerical procedures construct a sample using $p(\theta^T|V)$ defined by (2.13). This appendix verifies that these procedures are consistent with the non-linear transition function defined in the text. In particular, we verify the nonlinear transition equation, $p(\theta_{t+1}|\theta_t, V) \propto I(\theta_{t+1})f(\theta_{t+1}|\theta_t, V)$ from equations (2.3), (2.13), (2.14), and (2.15). First consider the transition equation for terminal state,

$$p(\theta_T|\theta_{T-1}, V) = \frac{p(\theta_T, \theta_{T-1}|V)}{p(\theta_{T-1}|V)}. \tag{7.1}$$

The joint density in the numerator can be expressed as

$$\begin{aligned} p(\theta_T, \theta_{T-1}|V) &= \int p(\theta^T|V)d\theta^{T-2} \\ &\propto I(\theta_T)f(\theta_T|\theta_{T-1}, V) \int \prod_{t=0}^{T-2} I(\theta_{t+1})f(\theta_{t+1}|\theta_t, V)d\theta^{T-2}. \end{aligned} \tag{7.2}$$

[32] Prospects for a gradual backsliding away from the zero inflation Ramsey outcome toward the higher Nash inflation rate also permeate the 'mean dynamics' in the model of Sargent (1999) and Cho, Williams, and Sargent (2001).

The marginal density in the denominator of (7.1) can be expressed as

$$\begin{aligned} p(\theta_{T-1}|V) &= \int p(\theta_T, \theta_{T-1}|V) d\theta_T \qquad (7.3) \\ &\propto \int I(\theta_T) f(\theta_T|\theta_{T-1}, V) d\theta_T \int \prod_{t=0}^{T-2} I(\theta_{t+1}) f(\theta_{t+1}|\theta_t, V) d\theta^{T-2}. \end{aligned}$$

The ratio between the two is

$$p(\theta_T|\theta_{T-1}, V) \propto I(\theta_T) f(\theta_T|\theta_{T-1}, V). \qquad (7.4)$$

Next consider the transition equation for the penultimate state,

$$p(\theta_{T-1}|\theta_{T-2}, V) = \frac{p(\theta_{T-1}, \theta_{T-2}|V)}{p(\theta_{T-2}|V)}. \qquad (7.5)$$

The joint density in the numerator of (7.5) can be expressed as

$$\begin{aligned} p(\theta_{T-1}, \theta_{T-2}|V) &= \int \int p(\theta^{T-1}|V) p(\theta_T|\theta_{T-1}, V) d\theta^{T-3} d\theta_T \qquad (7.6) \\ &= \int p(\theta^{T-1}|V) d\theta^{T-3} \int p(\theta_T|\theta_{T-1}, V) d\theta_T \\ &= \int p(\theta^{T-1}|V) d\theta^{T-3} \end{aligned}$$

where the last equality follows from the fact that $p(\theta_T|\theta_{T-1}, V)$ integrates to one. Using the same argument as above, this can be expressed as

$$p(\theta_{T-1}, \theta_{T-2}|V) \propto I(\theta_{T-1}) f(\theta_{T-1}|\theta_{T-2}, V) \int \prod_{t=0}^{T-3} I(\theta_{t+1}) f(\theta_{t+1}|\theta_t, V) d\theta^{T-3}.$$

The marginal density for θ_{T-2} is

$$\begin{aligned} p(\theta_{T-2}|V) &= \int p(\theta_{T-1}, \theta_{T-2}|V) d\theta_{T-1} \qquad (7.7) \\ &\propto \int I(\theta_{T-1}) f(\theta_{T-1}|\theta_{T-2}, V) d\theta_{T-1} \int \prod_{t=0}^{T-3} I(\theta_{t+1}) f(\theta_{t+1}|\theta_t, V) d\theta^{T-3}. \end{aligned}$$

The ratio between the two is

$$p(\theta_{T-1}|\theta_{T-2}, V) \propto I(\theta_{T-1}) f(\theta_{T-1}|\theta_{T-2}, V). \qquad (7.8)$$

Continuing a backward recursion implies

$$p(\theta_t|\theta_{t-1}, V) \propto I(\theta_t) f(\theta_t|\theta_{t-1}, V). \qquad (7.9)$$

Hence, the nonlinear transition equation can indeed be expressed in terms of the truncated linear transition equation.

8. References

Ahmed, Shagil and John H. Rogers, 1998, "Inflation and the Great Ratios: Long-term Evidence from the U.S." International Finance Discussion Paper No. 628, Board of Governors of the Federal Reserve System.

Albanesi, Stephania, V.V. Chari, and Lawrence J. Christiano, 2000, "Expectations Traps and Monetary Policy." Mimeo, Northwestern University and University of Minnesota.

Anderson, Brian D.O. and John B. Moore, 1979, *Optimal Filtering*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey).

Barro, Robert J. and David B. Gordon, 1983, "A Positive Theory of Monetary Policy in a Natural Rate Model." *Journal of Political Economy* 91, pp. 589-610.

Bernanke, Ben S. and Ilian Mihov, 1998a, "The Liquidity Effect and Long-Run Neutrality." In *Carnegie-Rochester Conference Series on Public Policy*, 49, Bennett T. McCallum and Charles I. Plosser, eds. (Amsterdam: North Holland), pp. 149-194.

Bernanke, Ben S. and Ilian Mihov, 1998b, "Measuring Monetary Policy." *Quarterly Journal of Economics*, 113, August, pp. 869-902.

Beveridge, Stephen and Charles R. Nelson, 1981, "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle,"' *Journal of Monetary Economics* 7, pp. 151-174.

Cho, In Koo, Noah Williams, and Thomas J. Sargent, 2001, "Escaping Nash Inflation," *Review of Economic Studies*, in press.

Chung, Heetaik, 1990, "Did Policy Makers Really Believe in the Phillips Curve? An Econometric Test," Ph.D. Dissertation, University of Minnesota, November.

Clarida, Richard, Jordi Gali, and Mark Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics* 115(1), pp. 147-180.

De Long, J. Bradford, 1997, "America's Only Peacetime Inflation: The 1970s," in *Reducing Inflation: Motivation and Strategy*, edited by Christina D. Romer and David Romer, NBER Studies in Business Cycles, vol. 30 (University of Chicago Press: Chicago, IL).

Doan, Thomas, Robert Litterman, and Christopher Sims, 1984, "Forecasting and Conditional Projections Using Realistic Prior Distributions," *Econometric Reviews* 3(1), pp. 1-100.

Eisner, Robert, 1997, "A New View of the NAIRU," in *Improving the Global Economy: Keynesianism and Growth in Output and Employment*, edit by Paul Davidson and Jan Kregel (Edward Elgar: Cheltenham).

Estrella, Arturo and Frederic S. Mishkin, 1999, "Rethinking the Role of the NAIRU in Monetary Policy: Implications of Model Formulation and Uncertainty," in *Monetary Policy Rules*, edited by John B. Taylor, NBER Conference Report (University of Chicago Press: Chicago, IL)

Fair, Ray, 1996, "Testing the Standard View of the Long-Run Unemployment-Inflation Relation," Cowles Commission Discussion Paper 1121.

Fisher, Mark E. and John J. Seater, 1993, "Long-Run Neutrality and Superneutrality in an ARIMA Framework," *American Economic Review* 83(3), pp. 402-415.

Friedman, Benjamin, 1998, "Introduction," in *Inflation, Unemployment, and Monetary Policy*, edited by Robert M. Solow and John B. Taylor (MIT Press: Cambridge, MA).

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin, 1995, *Bayesian Data Analysis* (Chapman and Hall, London).

Gordon, Robert J., 1970, "The Recent Acceleration of Inflation and Its Lessons for the Future," *Brookings Papers on Economic Activity,* No. 1, pp. 8-41.

Granger, Clive W.J., 1966, "The Typical Spectral Shape of an Economic Variable," *Econometrica* 34(1), pp. 150-161.

Hall, Robert, 1999, "Comment on Rethinking the Role of the NAIRU in Monetary Policy: Implications of Model Formulation and Uncertainty," in *Monetary Policy Rules*, edited by John B. Taylor, NBER Conference Report (University of Chicago Press, Chicago, Illinois)

Ireland, Peter, 1999, "Does the Time-Consistency Problem Explain the Behavior of Inflation in the United States?" *Journal of Monetary Economics* 44(2), pp. 279-292.

Kim, Chang-Jin and Charles R. Nelson, 1999, *State-Space Models with Regime Switching* (MIT Press: Cambridge, Massachusetts).

King, Robert G. and Mark W. Watson, 1994, "The Post-War U.S. Phillips Curve: A Revisionist Econometric History," *Carnegie-Rochester Conference Series on Public Policy*, vol. 41, pp. 157-219.

___ and ___, 1997, "Testing Long-Run Neutrality," Federal Reserve Bank of Richmond *Economic Quarterly* 83(3), pp. 69-101.

Kydland, Fynn and Edward C. Prescott, 1977, "Rules Rather than Discretion: the Inconsistency of Optimal Plans," *Journal of Political Economy* 85 (3), pp. 473-491.

Lucas, Robert E., Jr., 1972, "Econometric Testing of the Natural Rate Hypothesis,: in *The Econometrics of Price Determination*, edited by Otto Eckstein (Board of Governors of the Federal Reserve System: Washington, DC).

___, 1976, "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and Labor Markets*, edited by Karl Brunner and Alan Meltzer, Carnegie-Rochester Series on Public Policy, vol. 1.

McCallum, Bennett T. and Edward Nelson, "Performance of Operational Policy Rules in an Estimated Semiclassical Structural Model," in *Monetary Policy Rules*, edited by John B. Taylor, University of Chicago Press, pp. 15-45.

Parkin, Michael, 1993, "Inflation in North America," in *Price Stabilization in the 1990s*, edited by Kumiharo Shigehara.

Phelps, Edmund S. 1967, "Phillips Curves, Expectations of Inflation, and Optimal Unemployment Over Time," *Economica* 2(3), pp. 22-44.

Roberts, G.O. and A.F.M. Smith, 1992, "Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms," *Stochastic Processes and their Applications* 49, pp. 207-216.

Rudebusch, Glenn D. and Lars E.O. Svensson, 1999, "Policy Rules for Inflation Targeting," in *Monetary Policy Rules*, edited by John B. Taylor, NBER Conference Report (University of Chicago Press, Chicago, Illinois)

Sargent, Thomas J., 1971, "A Note on the Accelerationist Controversy," *Journal of Money Credit and Banking* 8(3), pp. 721-725.

___, 1983, "Autoregressions, Expectations, and Advice, *American Economic Review, Papers and Proceedings*, Vol. 74, No. 2, pp. 408-415.

___, 1999, *The Conquest of American Inflation* (Princeton University Press: Princeton, New Jersey).

Sims, Christopher A., 1982, " Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity*, Vol. 1, pp. 107-152.

Sims, Christopher A., 1988, "Projecting Policy Effects with Statistical Models," *Revista de Analysis Economico* 3, pp. 3-20.

Sims, Christopher A., 1999. "Drifts and Breaks in Monetary Policy," mimeo, Princeton University.

Solow, Robert M., 1968, "Recent Controversy on the Theory of Inflation: An Eclectic View," in *Proceedings of a Symposium on Inflation: Its Causes, Consequences, and Control*, edited by Stephen Rousseaus (New York University: New York).

___,1998, "How Cautious Must the Fed Be?" in *Inflation, Unemployment, and Monetary Policy*, edited by Robert M. Solow and John B. Taylor (MIT Press: Cambridge, MA).

Taylor, John B., 1997, "Comment on America's Only Peacetime Inflation: The 1970s," in *Reducing Inflation: Motivation and Strategy*, edited by Christina D. Romer and David Romer, NBER Studies in Business Cycles, vol. 30 (University of Chicago Press: Chicago, IL).

___, 1998, "Monetary Policy Guidelines for Unemployment and Inflation Stability," in *Inflation, Unemployment, and Monetary Policy*, edited by Robert M. Solow and John B. Taylor (MIT Press: Cambridge, MA).

___, 1999, "Staggered Price and Wage Setting in Macroeconomics," in Handbook of Macroeconomics, vol. 1B, edited by John B. Taylor and Michael Woodford (North Holland, Amsterdam).

Tobin, James, 1968, "Discussion," in *Proceedings of a Symposium on Inflation: Its Causes, Consequences, and Control*, edited by Stephen Rousseaus (New York University: New York).

Whittle, Peter, 1953, "The Analysis of Multiple Stationary Time Series," *Journal of the Royal Statistical Society*, Series B, vol. 15, pp. 125-139.