Tractable Bayesian Optimal Policy in the Presence of Regime Change and Local Parameter Uncertainty

Eric T. Swanson
Federal Reserve Bank of San Francisco
eric.swanson@sf.frb.org
http://www.ericswanson.pro

Abstract

This paper proposes an approximation to the optimal policy problem in forward-looking models with regime change that allows for tractable solution for the optimal policy even when the parameters of the model and the current regime are not known with certainty. The linear-quadratic framework is adhered to as much as possible, with a particular emphasis on maintaining the property of separation of estimation and control, which is crucial for maintaining tractability. Generality is achieved by allowing for full Bayesian updating of all aspects of the model, including model parameters, the probability that a regime change has occurred, and the values of all (generally non-normal) shocks hitting the model each period. The methods of this paper provide results that can be quite useful in practice—for example, they fit the policy behavior of the Federal Reserve in the late 1990s very well, which standard methods fail to do.

JEL Classification: E52

Version 0.2 — still very preliminary and incomplete

June 13, 2006

1. Introduction

Since the rational expectations revolution of the 1980s, there has been tremendous interest in computing the optimal policy in forward-looking economic models. Nevertheless, optimal policy in these models is very difficult to compute in general, particularly when one recognizes and attempts to incorporate the vast variety of uncertainty that is faced by policymakers in practice. For example, policymakers typically face uncertainty about the values of the parameters in the economic model, uncertainty about *which* among several candidate models provides the best description of the true economic system that they wish to control, and uncertainty associated with the possibility that the parameters or relationships in the model may undergo periodic structural breaks (or regime changes) over time.

The interaction of this uncertainty with the forward-looking behavior of the model gives rise to an enormous curse of dimensionality in general, since the policymaker's choice of the policy instrument today significantly affects the policymaker's distribution of beliefs tomorrow via learning, which in turn feeds back to the optimal choice of the policy instrument today. Thus, the policymaker's entire distribution of beliefs about unobserved elements of the model must effectively be treated as an (infinite-dimensional) state vector in the policymaker's optimal control problem.

As a result, researchers in the area have found it essential to find simplifications and approximations to the problem that allow for tractability while still providing insight into the most important characterisics of the optimal policy in the model. For example, in Cogley, Colacito, and Sargent (2005), the dimensionality of the policymaker's beliefs are drastically reduced by assuming that there are only two possible models, each with *known* parameter values, and one of which must be true.[1] Thus, the policymaker's beliefs at any point in time can be completely described by a single real number—the probability that model A is true as opposed to model B—which keeps the state space of the optimal control problem small and makes the problem tractable. Zampolli (2004), Blake and Zampolli (2005), and Svensson and Williams (2005) consider a different approximation,

[1] The authors also allow for the possibility that a third model generates the data, but this does not change the essntial point that policymaker's beliefs are reduced to a single degree of freedom.

adapting the literature on Markov-jumping linear-quadratic models from the engineering literature to the forward-looking case to allow for multiple linear-quadratic regimes with occasional Markovian transitions between them. The key assumptions in this case are that the parameters of each model and *which* model is in force today are all known by the policymaker with certainty, which preserves the main features of linear-quadratic optimal control and again maintains tractability of the problem.

The present paper proposes a new and alternative approximation to the optimal policy problem in the presence of regime change and parameter uncertainty that allows for tractable solution for the optimal policy even when the parameters of the model and the current regime are not known with certainty. As in Zampolli (2004), Blake and Zampolli (2005), and Svensson and Williams (2005), the linear-quadratic framework is adhered to as much as possible, with a particular emphasis on maintaining the property of *separation of estimation and control,* which is an indispensible tool for maintaining tractability of the problem. Generality is achieved by allowing for full Bayesian updating of all parameters, the probability that a regime change has occurred, and all (generally non-normal) shocks each period. Although Bayesian updating of these general distributions is computationally nontrivial, it is nonetheless *much* more tractable than attempting to solve problems that violate the separation of estimation and control, for the reasons discussed previously. The methods of this paper also give results that seem to be quite useful in practice—for example, Swanson (2006) showed that a less general version of the framework in this paper fits the Federal Reserve's policy in the late 1990s extremely well.

The methods in this paper build most directly on previous work by Pearlman (1992), Svensson and Woodford (2003), Swanson (2004), Meyer, Swanson, and Wieland (2001), and Swanson (2006). Those papers derive and analyze the implications of optimal policy in the linear-quadratic framework with imperfect observation of the state variables of the model. The present paper generalizes these earlier analyses by showing how to incorporate regime change and local parameter uncertainty into the framework.

The remainder of the paper proceeds as follows. Section 2 defines the model with regime change and no parameter uncertainty. Section 3 shows how to solve this model and presents an example of the U.S. in the late 1990s. Section 4 introduces local parameter uncertainty into the model and shows how to solve this case. Section 5 provides a detailed example that works through the solution algorithm and carries along the distribution of the policymaker's beliefs over time. Section 6 concludes.

2. A Model with Regime Change

Before turning to the full model with regime change and parameter uncertainty in section 4, we first analyze the optimal monetary policy in a model with regime change and no parameter uncertainty.

We assume that the economy is described by an $n \times 1$ system of linear equations:[2]

$$y_t = Ay_{t-1} + BE_t y_{t+1} + Cx_t + Dr_t + \varepsilon_t, \tag{1}$$

where y_t denotes an $n \times 1$ vector of endogenous variables, r_t an $n_r \times 1$ vector of control variables, x_t an $n_x \times 1$ vector of variables that are exogenous with respect to y and r, ε_t denotes an $n \times 1$ vector of stochastic shocks, and E_t the mathematical expectation operator conditional on all information at time t (which we will specify shortly). Intuitively, one may think of y_t as representing endogenous variables such as inflation, GDP, and the output gap, and x_t as representing exogenous variables such as potential output, the natural rate of unemployment, the natural rate of productivity growth, and the natural (real) rate of interest. The shocks ε_t are assumed to be independent over time with mean zero, distribution function $F_{\varepsilon t}$ which may vary over time t, and variance matrix $\Omega_{\varepsilon t}$. The (exogenous) stochastic process for the variables x_t will be specified shortly.

In each period t, a policymaker sets the control vector r_t to minimize an expected discounted quadratic loss:

$$(1-\delta)E_t \sum_{s=t}^{\infty} \delta^s z_s' Q z_s, \tag{2}$$

where $z_t' \equiv [x_t', y_t', r_t']$ and $\delta \in (0,1)$. We will consider the optimal policy under both commitment and discretion below.

[2] The model in (1) may be rewritten in terms of "predetermined" and "nonpredetermined" endogenous variables if that transformation is preferred. The Anderson-Moore (1985) algorithm provides such a transformation and Dennis (2006) provides additional discussion of the merits of each representation.

In forming expectations at time t, the policymaker knows all the parameters of the model (e.g., A, B, C, D, F, and $\Omega_{\varepsilon t}$) and the current and past values of the control vector r_s, $s \leq t$, but does not necessarily observe all the elements of x_s and y_s, $s \leq t$. Instead, the policymaker observes the history of an $n_w \times 1$ vector of indicator variables, which may or may not include elements of x and y:

$$w_s \equiv G z_s + \eta_s, \quad s \leq t, \tag{3}$$

where η_t is an $n_w \times 1$ vector of stochastic shocks that is *i.i.d.* over time with mean zero, distribution function $F_\eta(\cdot)$, and variance Ω_η. Note that by setting the appropriate elements of G and Ω_η to zero, the control vector r_s can be included within the vector of indicator variables w_s.

Finally, the exogenous variables x_t and the stochastic shocks ε_t are potentially subject to regime change. (There is no conceptual difficulty in allowing the endogenous variables y to undergo regime change as well, but in practice this case tends to be less interesting.) Regime change follows an exogenous Poisson arrival process with probability p_x of occuring in each period t. We define the random variable χ_t to equal 1 if a regime change occurs in period t, and 0 otherwise. The policymaker does not observe the arrival of a regime change directly, but must instead infer it from the behavior of the observable variables w. Thus, the policymaker is essentially never absolutely sure which regime he or she is actually in.

The variables x are affected by regime change in an additive way. For example, if the natural rate of unemployment u_t^* is an element of x_t, we might have:

$$u_t^* = \begin{cases} u_{t-1}^* & \text{if } \chi_t = 0, \\ u_{t-1}^* + \Psi_t & \text{if } \chi_t = 1, \end{cases}$$

where Ψ_t is a random variable that denotes the way u_t^* jumps in the event that a regime change occurs. Intuitively, one can think of the probability of regime change p_x as being small (e.g., 1 percent per quarter), so that regime change is a relatively infrequent event.

More generally, we will allow for the case where any of the exogenous variables in x may or may not be subject to stochastic drift and to regime change each period, with

$$x_t = \begin{cases} H x_{t-1} + \nu_t & \text{if } \chi_t = 0, \\ H x_{t-1} + \nu_t + \Psi_t & \text{if } \chi_t = 1, \end{cases} \tag{4}$$

where ν_t is an $n_x \times 1$ vector of stochastic shocks that is *i.i.d.* over time with mean zero, distribution function $F_\nu(\cdot)$, and variance Ω_ν, and Ψ_t is an $n_x \times 1$ stochastic shock vector with mean zero, distribution function $F_\Psi(\cdot)$, and variance Ω_Ψ. Intuitively, one can think of the elements of Ω_Ψ as being much larger than the elements of Ω_ν. Note that one or more elements of x may be subject to regime change within this framework, and those elements may be subject to regime change separately or simultaneously, depending on how the distribution function F_Ψ is specified. There is also no requirement that Ω_Ψ be diagonal; thus, when a regime change occurs, it can potentially affect all elements of x_t in a way that may be related (e.g., the natural rate of productivity growth and the natural rate of interest), as specified by F_Ψ.

In addition, the distribution function $F_{\varepsilon t}$ for the shocks ε_t is potentially subject to regime change in each period. Regime change in $F_{\varepsilon t}$ is independent of regime change in x, and follows a Poisson arrival process with arrival probability p_ε each period. We let the distribution function $F_{\varepsilon t}$ at time t be indexed by the variable θ_t, so that $F_{\varepsilon t}(\cdot) = F_\varepsilon(\cdot, \theta_t)$. Regime change in $F_{\varepsilon t}$ can thus be indexed by a change in θ_t. We let the random variable φ_t take on the value 1 when a regime change in θ_t occurs, and the value 0 otherwise. In particular, we define:

$$\theta_t = \begin{cases} \theta_{t-1} & \text{if } \varphi_t = 0, \\ \theta_{t-1} + \Phi_t & \text{if } \varphi_t = 1, \end{cases} \tag{5}$$

where Φ_t is a scalar stochastic shock with mean zero, distribution function $F_\Phi(\cdot)$, and variance Ω_Φ. Thus, we can allow the variances of the shocks in the model to change when a regime change occurs, as has been suggested was the case for the U.S. in the 1980s by McConnell and Perez-Quiros (2000), Stock and Watson (2002), Sims and Zha (2005), and others.

Shocks dated t in the model are independent of all variables dated $t-1$ and earlier. For simplicity, we assume that ε_t, η_t, ν_t, Ψ_t, and Φ_t are independent of each other at time t. It is straightforward to extend the model to the case where the distribution functions F_i and variances Ω_i, $i \in \{\varepsilon, \eta, \nu, \Psi, \Phi\}$, vary over time, so long as that variation is exogenous with respect to the control vector r and endogenous variables y.

The policymaker's information set when forming expectations at each date t is thus:

$$\mathcal{I}_t \equiv \{A, B, C, D, F_\varepsilon, F_\eta, F_\nu, F_\Psi, F_\Phi, G, H, p, w_s | s \leq t\}, \tag{6}$$

together with a time-0 prior distribution over all the variables of the model. The variances Ω_i, $i \in \{\eta, \nu, \Psi, \Phi\}$, are also known by the policymaker at time t but are not explicitly listed in (6) because they are implied by the distribution functions $F_i(\cdot)$. If $\Omega_{\varepsilon t}$ is time-varying, then $\Omega_{\varepsilon t}$ is typically not known by the policymaker with certainty because φ_t is typically unobserved. Recall that the history of the control vector r is also known by the policymaker and is included in the history of observable variables w. Finally, note that we have adopted the standard timing convention in forward-looking models with imperfect information that the control vector r_t, expectations E_t, and all variables dated t (y_t, w_t, x_t, ε_t, η_t, ν_t, Ψ_t, and Φ_t) are determined simultaneously and jointly at time t. Svensson and Woodford (2003) and Woodford (2003) discuss this simultaneity (or "circularity" in their terminology) in some detail, while Pearlman, Currie, and Levine (1986) provide an early example and discussion.

Finally, recall that model (1)–(6) abstracts away from parameter uncertainty and from regime change in the parameters of the model (other than possible regime change in the variances of the shocks). Thus, in this section we are implicitly focusing attention on questions of the form, "Has the natural rate of unemployment or productivity growth changed?" or "Have the variances of the shocks ε changed?", as opposed to questions of the form "Has the slope of the Phillips curve or the IS curve changed?" Of course, which of these questions is of primary concern for policymakers is a matter that may be subject to debate, but the experience of the U.S. in the mid-to-late 1990s illustrates that questions of the first type are indeed extremely relevant for the conduct of optimal monetary policy in practice.

3. Optimal Policy for the Model with Regime Change

The presence of regime change and the possibility that shocks are non-normally distributed in the model (1)–(6) leads to a complicated nonlinear filtering problem for policymakers. At each date t, policymakers must update their beliefs about $\{x_s, y_s, \varepsilon_s, \eta_s, \nu_s, \Psi_s, \Phi_s\}$, $s \leq t$, via Bayesian updating based on the history of observable variables w_s, $s \leq t$.

As mentioned in the Introduction, the interaction of the optimal policy problem in a forward-looking model with the nonlinear filtering problem is computationally intractable in general. However, a few key properties of model (1)–(6) render the solution to that model computationally tractable, despite its generality and applicability. In particular, equations (1)–(6) form a system of linear equations with quadratic objective, and we have carefully specified the definition of regime change in the model to preserve this underlying structure. For example, the state variable θ_t can be incorporated into the vector x_t without loss of generality, and the shock Φ_t can be incorporated into the shock Ψ_t.

In this way, regime change in the model, despite being applicable to a wide variety of cases that have been studied in the empirical literature, takes the form of a simple additive (though non-normally distributed) shock $\tilde{\nu}_t \equiv \nu_t + \Psi_t$.

The additive nature of the uncertainty in the model—both the regime change shock Ψ_t and the more standard shocks ε_t, η_t, and ν_t—preserves the linear-quadratic structure of the model despite the possible non-normality of all of the shocks (and the definite non-normality induced by the possibility of changes in regime). As a result, the basic insights of Pearlman (1992) and Svensson and Woodford (2003) regarding optimal policy in forward-looking linear-quadratic models can be applied. The most important of these insights is stated in the following two propositions:

Proposition 1: *The optimal policy without commitment in model (1)–(6) exhibits the property of separation of estimation and control. In particular, there exist matrices J, K, L, M, and N such that the optimal choice for r_t at each time t satisfies:*

$$r_t = JE_t y_{t-1} + KE_t x_{t-1} + LE_t \varepsilon_t + ME_t \nu_t + NE_t \Psi_t. \tag{7}$$

Proposition 2: *The optimal policy with commitment in model (1)–(6) exhibits the property of separation of estimation and control. In particular, there exist matrices J, K, L, M, N, and P such that the optimal choice for r_t at each time t satisfies:*

$$r_t = JE_t y_{t-1} + KE_t x_{t-1} + LE_t \varepsilon_t + ME_t \nu_t + NE_t \Psi_t + P\xi_{t-1}, \tag{8}$$

where ξ_t denotes the vector of Lagrange multipliers associated with the endogenous variables of the model.

Proof of propositions 1 and 2: See Pearlman (1992). Svensson and Woodford (2003) provide alternative proofs and additional discussion. □

Note that model (1)–(6) and equations (7) and (8) can be rewritten in terms of "predetermined" and "nonpredetermined" endogenous variables if that transformation is preferred. Note that the conditional expectations of the shocks in equations (7) and (8) are not zero in general because of the standard timing assumption that expectations at time t are formed jointly and simultaneously with the setting of the control vector r and the realizations of the observable indicators w, the shocks ε, η, ν, and Ψ, and the endogenous and exogenous variables y and x.

The matrices J, K, L, M, and N in equations (7) and (8) in general differ across these two equations. In both equations, the solution matrices J through N (or P) must typically be computed numerically using straightforward generalizations of the Blanchard-Kahn (1980) algorithm (e.g., Anderson and Moore, 1985) applied to a policymaker Lagrangean or defining a value function and iterating on a matrix Riccatti equation. In the case of commitment (Proposition 2), the auxiliary Lagrange multipliers must be defined in terms of the policymaker's Lagrangean and capture the constraint that (state-contingent) future policy that is promised by the policymaker at date 0 must be upheld once those future periods arrive—additional discussion and references are provided in the papers cited above.

In this paper, we will not develop these solution procedures for the optimal policy matrices J through P any further, since these techniques are well understood and widely available from other sources. The essential insight here is that these techniques are applicable to the model of regime change given by equations (1)–(6). Thus, to solve for the optimal policy in that model, we may take it as given that the computation of the matrices J through P is straightforward and focus attention on the nonlinear filtering problem of the underlying state of the economy. While the policymaker's nonlinear filtering problem is nontrivial, once it is decoupled from the optimal control problem (the matrices J through P), it too becomes tractable using off-the-shelf Bayesian updating methods. The tractability of the procedure is perhaps best illustrated with a simple example, as follows.

3.1 An Example: The U.S. in the Late 1990s

For concreteness and simplicity, consider the case of a single unobserved exogenous variable, u_t^*, which one may think of as denoting the natural rate of unemployment. Assume that u^* is constant over time except for the possibility of a regime change, which follows the

Poisson arrival process described in the previous section. Thus, u_t^* satisfies:

$$u_t^* = u_{t-1}^* + \Psi_t, \tag{9}$$

$$\Psi_t = \begin{cases} 0 & \text{with probability } 1-p, \\ X \sim N(0,\sigma^2) & \text{with probability } \quad p. \end{cases}$$

Assume also that there is a single endogenous variable, $\tilde{u}_t$, which may be thouht of as the "unemployment gap" (the deviation of unemployment from its natural level), and which evolves according to:

$$\tilde{u}_t = \alpha\tilde{u}_{t-1} + \beta E_t\tilde{u}_{t+1} - \gamma r_t + \varepsilon_t, \tag{10}$$

where r_t can be thought of as the real interest rate, which is set by the policymaker.

Assume that there is a single indicator variable, u_t, which may be thought of as the observed unemployment rate, and which satisfies:

$$u_t = u_t^* + \tilde{u}_t + \eta_t. \tag{11}$$

Finally, policymakers seek to minimize a discounted sum of expected deviations of unemployment from its natural rate,

$$(1-\delta)E_t\sum_{s=t}^{\infty}\delta^{s-t}\tilde{u}_s. \tag{12}$$

The optimal policy under discretion in the above model then satisfies:

$$r_t = \theta E_t\tilde{u}_{t-1} + \varphi E_t\varepsilon_t, \tag{13}$$

for some coefficients θ and φ. The optimal policy under commitment satisfies:

$$r_t = \theta E_t\tilde{u}_{t-1} + \varphi E_t\varepsilon_t + \mu\xi_t, \tag{14}$$

for some coefficients θ, φ, and μ.

We initialize the model in period 0 with a prior distribution on u^* that is normally distributed with mean 6 and standard deviation 0.3, which is a rough calibration to the U.S. economy in the mid-1990s. We choose a normal distribution as the initial policymaker prior to demonstrate that even starting from a purely Gaussian prior, the presence of regime change in the model will naturally lead over time, through the policymaker's nonlinear filtering problem, to non-normal yet intuitive posterior distributions. Moreover, after many periods without a regime change, policymakers' beliefs would evolve to a distribution that is close to normally distributed anyway, so a normal distribution is also a rough calibration to the U.S. economy in the mid-1990s, when it is thought that there had not been a substantial structural break in u^* for many years. We set $p = .01$ and $\sigma = 1$, so that a break in u^* occurs roughly once every 25 years and each break draws a jump in u^* from a $N(0, 1)$ distribution. Based on quarterly U.S. data, we set $\alpha = 0.5$, $\beta = 0.3$, $\gamma = 0.1$, and $\sigma_\varepsilon = .3$ in equation (10),[3] and assume that $\delta = 0.99$. As in the previous section, these parameters are all known by policymakers with certainty.

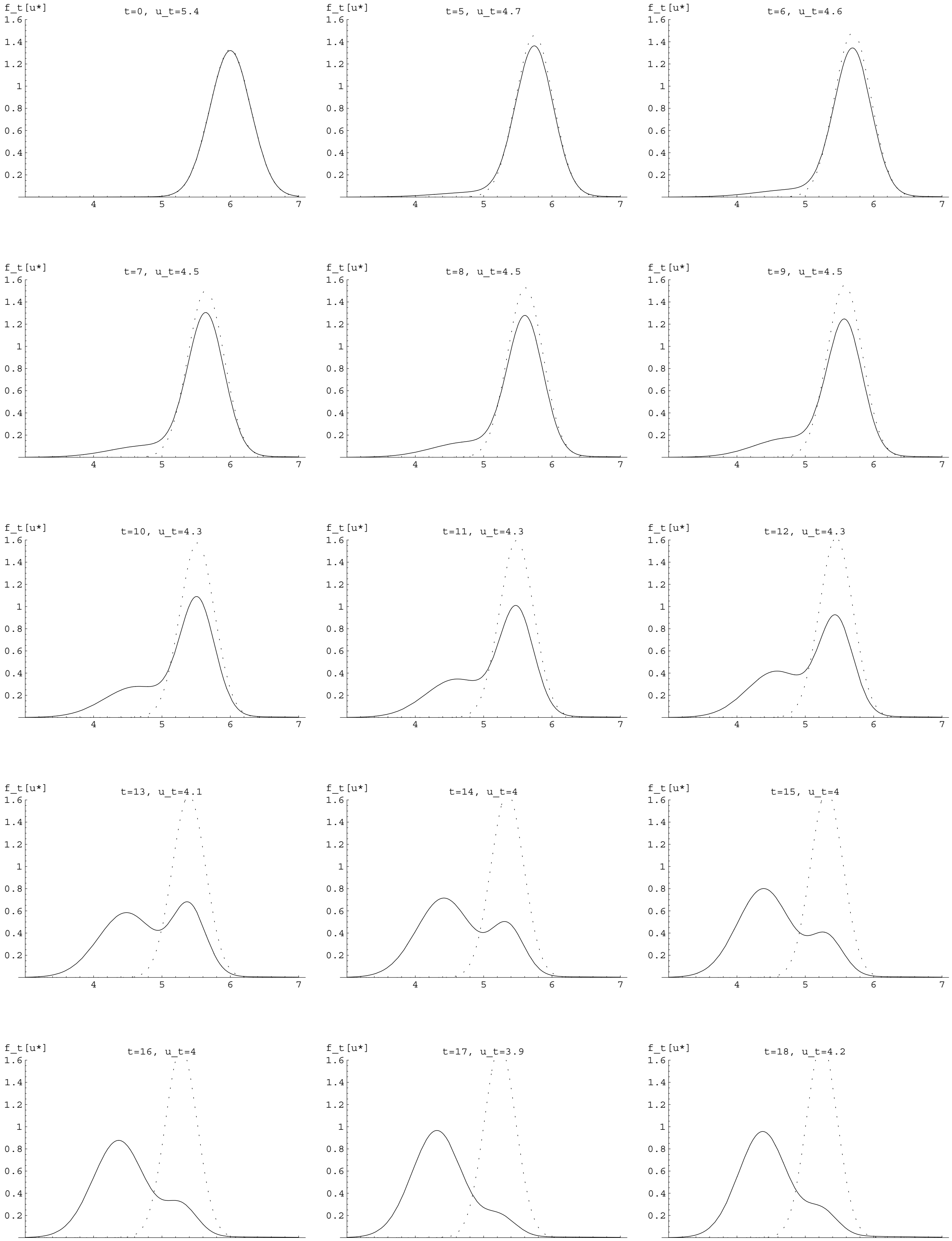
Figure 1 presents the results from simulating this model forward for 18 quarters. The solid line in each panel plots the policymaker's posterior distribution on u_t^* in each period t for the model described above, while the dotted line plots results for the same model *without* structural change (i.e., setting $p = 0$) for comparison. The time 0 prior for each model is plotted in the upper left-most panel; the results for the first four quarters ($t = 1, \ldots, 4$) are not reported in the figure because they are not very different from the $t = 0$ and $t = 5$ cases. Simulating the model forward requires drawing values for u_t each period; we use the quarterly values of unemployment realized in the U.S. from 1997Q1 through 2001Q1, which are listed at the top of each panel.[4]

The simulation becomes interesting after about 5 or 6 quarters. By this point, repeated realizations of u_t below the policymaker's prior mean of 6 cause them to place noticeable probability on a structural break having occurred. This is seen as a fat left tail of the distribution for u^*. After about 11 or 12 quarters, the policymaker's beliefs display a clear bimodal pattern, with "New Economy" (low u^*) and "Old Economy" (high u^*) regimes. Once the policymaker's beliefs reach this point, the policymaker is extremely uncertain about the true value of u^* over a wide range of intermediate values, and begins to revise beliefs very strongly as new information comes in. This uncertainty persists for

[3] We estimate these parameters by maximum likelihood over the period 1973–1996, using a simple HP-filter applied to u_t to measure $u_t - u^*$.

[4] Realized unemployment rates for $t = 1, \ldots, 4$ are 5.2, 5.2, 5.1, and 4.9, respectively.

Figure 1: Dynamic Simulation of Policymaker's Beliefs



several quarters, and is still evident 18 quarters from the start of the simulation in the bottom right panel.

By contrast, in the model without structural change (dotted lines in Figure 1), the policymaker's beliefs have a Gaussian distribution in every period. Even though the policymaker updates his or her beliefs each period in response to incoming data, the Gaussian prior and Gaussian shocks ensure that only the mean of the distribution is revised (with the variance gradually shrinking over time), while the overall Gaussian functional form is preserved.

The diffuseness of the distributions in the middle and lower panels of Figure 1 that is induced by the possibility of regime change leads the optimal policy to respond cautiously at first to surprises in the observed unemployment rate but to respond increasingly aggressively at the margin as the these surprises become cumulatively larger (Swanson (2006) provides additional discussion and proofs of this feature of the optimal policy). In Figure 2, we show that these features are also well matched by the actual behavior of the Fed in the late 1990s. The left-hand panel plots the quarterly U.S. unemployment rate from 1986 to 2001, along with a dashed horizontal line at $u_t = 6$ to depict a common estimate of the natural unemployment rate in the mid-1990s (e.g., Staiger, Stock, and Watson (1997)). The right-hand panel compares the Federal Reserve's actual interest rate policy (solid line) to what would have been prescribed by a Taylor-type rule (dashed line) based on GDP deflator inflation and the unemployment rate data in Figure 2a.[5]

The Taylor-type rule in Figure 2b tracks the Fed's actual behavior quite well until about 1997, at which point the two diverge noticeably. In particular, the Fed appears to have been much less aggressive between 1997 and 1999 than a Taylor rule based on previous estimates of u^*, or an optimal filter of u^* that failed to allow for the possibility of regime change, would have implied. Beginning sometime in 1999, however, the Fed appears to have returned to the marginal responsiveness implied by the Taylor Rule—i.e., the marginal prescriptions from the Taylor rule seem to coincide with the Fed's marginal interest rate moves in 1999 and 2000, even though the levels had diverged as a result of

[5] The Taylor-type rule in Figure 2b is given by $i_t = r^* + \pi_t + \alpha(\pi_t - \pi^*) - \beta(u_t - u^*)$, where $r^* = 2.75$, $\pi^* = 2$, $u^* = 6$, $\alpha = 0.5$, and $\beta = 1.8$, u_t denotes the unemployment rate in Figure 2a, and π_t the four-quarter change in GDP deflator inflation.

Figure 2a

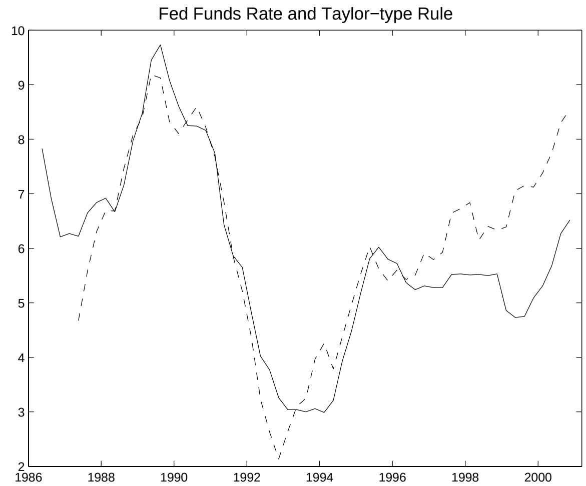


Figure 2b

the Fed's earlier cautiousness. The Fed's behavior over this episode—an attenuation in the responsiveness of policy for a time, followed by a return to a more aggressive response at the margin—matches very well the prescriptions for optimal policy that are implied by the dynamic simulation of the policymaker's beliefs in Figure 1.

3.2 Discussion

While the simulation above restricted attention to a univariate model for ease of exposition and to keep the figures clear and simple, the computation involved in keeping track of the policymaker's beliefs at each date t is quite tractable for relatively low (3 or 4) degrees of freedom of uncertainty and relatively large numbers of indicator variables. Thus, so long as the optimal policy problem can be framed in terms of a relatively small number of unobserved variables, tracking the policymaker's beliefs over time and thus computing the optimal monetary policy at any given date t does not pose an insurmountable computational challenge.

4. The Model with Local Parameter Uncertainty

We now turn to the case of parameter uncertainty. For ease of exposition, we abstract away from the possibility of regime change in this section, but in the next section we will work through the solution to a model in which both regime change and local parameter uncertainty are present.

As before, assume that the economy is described by an $n \times 1$ system of linear equations:

$$y_t = Ay_{t-1} + BE_t y_{t+1} + Cx_t + Dr_t + \varepsilon_t, \tag{15}$$

where now the parameters A, B, C, and D are not known with certainty by the policymaker but are instead uncertain. Let $\bar{A}$ denote the mean of policymaker's prior (time 0) distribution for A, $\bar{B}$ policymaker's prior mean for B, and so on, and let $dA \equiv A - \bar{A}$, $dB \equiv B - \bar{B}$, etc.

We can then write equation (15) as:

$$\begin{aligned} y_t = {} & \bar{A}y_{t-1} + \bar{B}E_t y_{t+1} + \bar{C}x_t + \bar{D}r_t + \varepsilon_t \\ & + dAy_{t-1} + dBE_t y_{t+1} + dCx_t + dDr_t. \end{aligned} \tag{16}$$

Let the policymaker's loss function and indicator variables of the model be defined as previously. For simplicity, assume that the parameter values in these equations, and variances of all shocks in the model, are known by the policymaker with certainty.

In general, the solution to the optimal policy problem with parameter uncertainty in even completely backward-looking models is extremely computationally intensive (see, e.g., Wieland, 2000a, 2000b, Beck and Wieland, 2002). However, there is an important special case to which we now turn.

Suppose first that the uncertainty surrounding the parameter matrices A, B, C, and D is small, in a sense that we will make more precise shortly. Second, suppose that the system of equations given by (16) is not the true, structural set of equations that describe the economy, but rather is a linearization (or log-linearization) of a nonlinear economic model around a nonstochastic steady state, as is standard in the literature. This second condition in particular requires that the exogenous, endogenous, and control variables of the model, x, y, and r must remain close to their steady-state values, $\bar{x}$, $\bar{y}$, and $\bar{r}$, in order for the system (16) to remain valid. Thus, the terms "close" and "small" here must be taken to refer to arbitrarily small neighborhoods of $(\bar{x}, \bar{y}, \bar{r})$ and $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$, as in a standard perturbation analysis (e.g., Judd, 1999, Swanson, Anderson, and Levin, 2006).

Note that several authors, notably Judd (1999) and Woodford (2003), have emphasized that there are conditions that must be met for a linear-quadratic approximation to yield a valid first-order approximation to the true, nonlinear optimal policy for a nonlinear model. A common assumption that is sufficient to guarantee validity of the LQ approximation is that the nonstochastic steady state of the nonlinear model is free from first-order distortions (e.g., Rotemberg and Woodford, 1999, Erceg, Henderson, and Levin, 2000). In this paper, we take it as given that the linear model and quadratic objective in which the policymaker is interested (described by equations (15)) provides a characterization of optimal policy that the policymaker finds useful, but implicitly the policymaker must have made assumptions of this sort in the background.

Let $\bar{x}$, $\bar{y}$, and $\bar{r}$ denote the steady-state values of x, y, and r that underlie the above approximation, and let $dx_t \equiv x_t - \bar{x}$, $dy_t \equiv y_t - \bar{y}$, and $dr_t \equiv r_t - \bar{r}$. Then we can rewrite equation (16) as:

$$\begin{aligned} y_t = \bar{A}y_{t-1} + \bar{B}E_t y_{t+1} + \bar{C}x_t + \bar{D}r_t + \varepsilon_t \\ + dA\bar{y} + dB\bar{y} + dC\bar{x} + dD\bar{r} \\ + dA dy_{t-1} + dB E_t dy_{t+1} + dC dx_t + dD dr_t. \end{aligned} \tag{17}$$

The first key insight to make is that the third row of equation (17) consists only of second-order terms. Thus, under the assumption that the original linearization (or log-linearization) of the original nonlinear model was valid, it follows that the terms in the third row of (17) are negligible by comparison. In other words, if the system of linear equations given by (15) was valid in the first place—which seems to be an appropriate maintained hypothesis—then it follows that we can drop the third row from (17) and compute the optimal policy in the vastly simpler and more tractable model:

$$\begin{aligned} y_t = \bar{A}y_{t-1} + \bar{B}E_t y_{t+1} + \bar{C}x_t + \bar{D}r_t + \varepsilon_t \\ + dA\bar{y} + dB\bar{y} + dC\bar{x} + dD\bar{r}. \end{aligned} \tag{18}$$

The optimal monetary policy in (18) is every bit as accurate as the optimal monetary policy in (17), which is in turn equivalent to (15).[6]

[6] In equation (18), we are implicitly assuming that the steady state values $\bar{x}$, $\bar{y}$, and $\bar{r}$ are known by

The second crucial insight to make is that the uncertainty in the second row of (18) is additive, which implies that equation (18) can be mapped into the standard forward-looking linear-quadratic framework, as follows. The first row of (18) is, of course, a standard linear equation with no parameter uncertainty and thus fits into the standard linear-quadratic framework immediately. Now define:

$$\Phi_t = \begin{cases} dA\bar{y} + dB\bar{y} + dC\bar{x} + dD\bar{r}, & t = 0, \\ 0, & t > 0. \end{cases} \tag{19}$$

This definition emphasizes the fact that the policymaker's prior uncertainty regarding the parameters A, B, C, and D can be thought of as an unobserved "shock" that hits the economy at time 0, but not at any other time $t > 0$. It is crucial to note that the policymaker will continue to *learn* about this shock Φ_0 forever and will continually update his or her estimate of Φ_0 based on observations of the observable variables w as they come in. But the "shock" itself is not a recurring one.

Substituting (19) into model (18) then yields:

$$y_t = \bar{A}y_{t-1} + \bar{B}E_t y_{t+1} + \bar{C}x_t + \bar{D}r_t + \varepsilon_t + \Phi_t, \tag{20}$$

where now the optimal policy problem includes the task of keeping track of and estimating the unboserved value of Φ_0 over time. But the form of (20) makes it clear that this problem is in fact isomorphic to the one we have already shown how to solve in the previous section—namely, the optimal policy in the linear-quadratic framework with non-normal shocks and full Bayesian updating.

4.1 Discussion

The crucial assumption in this section that maintains the tractability of the problem is that the parameter uncertainty is local rather than large. While this assumption might sound restrictive at first pass, in practice the method often yields quite good numerical approximations over nonlocal regions of the parameter space (e.g., Judd, 1999, Swanson, Anderson,

---

the policymaker with certainty. It is easy to relax this assumption, however, by letting $\bar{x}$, $\bar{y}$, and $\bar{r}$ denote the policymaker's prior mean on the steady-state values of these variables. Under the assumption of local parameter uncertainty, the policymaker's uncertainty surrounding the true steady-state values of x, y, and r will also be local and thus can be dropped from the first-order approximation (18), just as we were able to drop the dA, dB, dC, and dD terms from (17).

and Levin, 2006), and local approximation methods are in fact a standard approach to the modeling of uncertainty in much previous research—e.g., Brock and Durlauf (2005), Cogley, Colacito, Hansen, and Sargent (2006), and the widespread use of log-linearization or perturbation methods more generally. Macroeconomic models in particular typically lend themselves very well to this kind of local approximation analysis, since the distance of the endogenous and exogenous variables y and x from steady state in these models tends to be quite small (amounting to no more than two or three percent at the outside) and the models themselves tend not to be extremely curved (except in those cases where risk aversion parameters are set in the hundreds or thousands).[7]

Thus, in practice the assumption of local parameter uncertainty provides a very useful theoretical framework for analysis of the problem without being unduly restrictive on the applicability of the method to standard macroeconomic applications. There is every reason to think that the methods of this section can be applied to the optimal policy problem in standard macroeconomic problems with nontrivial parameter uncertainty and yield excellent approximations to the true (nonlinear) optimal policy. We now turn to such an example.

5. Optimal Policy with Regime Change & Parameter Uncertainty

In this section we generalize the example from section 3 to consider the case of both regime change and local parameter uncertainty and work through the solution to this model, tracking the evolution of the policymaker's beliefs over time.

To be completed.

6. Conclusions

To be written.

[7] In support of this observation, Swanson, Anderson, and Levin (2006) show that the linear solution to a standard stochastic growth model in fact performs extremely well over large regions of the state space even when compared to solutions that are valid to much higher order.

References

Anderson, Gary and George Moore. "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models," *Economics Letters* 17 (1985), 247–52.

Beck, Günter and Volker Wieland. "Learning and Control in a Changing Economic Environment," *Journal of Economic Dynamics and Control* 26 (2002), 1359–77.

Blake, Andrew and Fabrizio Zampolli. "Time-Consistent Policy in Markov-Switching Models," unpublished manuscript, Bank of England (2005).

Brock, William and Steven Durlauf. "Local Robustness Analaysis: Theory and Application," *Journal of Economic Dynamics and Control* 29 (2005), 2067–92.

Cogley, Timothy, Ricardo Colacito and Thomas Sargent. "Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas," unpublished manuscript, New York University (2005).

Cogley, Timothy, Ricardo Colacito, Lars Hansen and Thomas Sargent. "Robustness and U.S. Monetary Policy Experimentation," unpublished manuscript, New York University (2006).

Dennis, Richard. "Optimal Policy in Rational Expectations Models: New Solution Agorithms," *Macroeconomic Dynamics* forthcoming (2006).

Erceg, Christopher, Dale Henderson, and Andrew Levin. "Optimal Monetary Policy with Staggered Wage and Price Setting," *Journal of Monetary Economics* 46 (2000), 281–313.

Hansen, Lars and Thomas Sargent. *Robustness* (Princeton: Princeton University Press, forthcoming 2006).

Judd, Kenneth. *Numerical Methods in Economics* (Cambridge, MA: MIT Press, 1999).

Meyer, Laurence, Eric Swanson, and Volker Wieland. "NAIRU Uncertainty and Nonlinear Policy Rules," *American Economic Review* 91 (2001), 226–31.

McConnell, Margaret and Gabriel Perez-Quiros. "Decomposing the Increased Stability of GDP Growth," *American Economic Review* 90 (2000), 1464–76.

Pearlman, Joseph. "Reputational and Nonreputational Policies under Partial Information," *Journal of Economic Dynamics and Control* 16 (1992), 339–57.

Pearlman, Joseph, David Currie, and Paul Levine. "Rational Expectations Models with Partial Information," *Economic Modelling* 3 (1986), 90-105.

Rotemberg, Julio and Michael Woodford. "Interest Rate Rules in an Estimated Sticky Price Model," in John Taylor, ed., *Monetary Policy Rules,* (Chicago: University of Chicago Press, 1999), 57–119.

Sims, Christopher and Tao Zha. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review,* forthcoming (2006).

Staiger, Douglas, James Stock, and Mark Watson. "How Precise Are Estimates of the Natural Rate of Unemployment?" in *Reducing Inflation: Motivation and Strategy,* ed. Christina Romer and David Romer (Chicago: Univ. of Chicago Press, 1997), 195–242.

Stock, James and Mark Watson. "Has the Business Cycle Changed and Why?" in Mark Gertler and Kenneth Rogoff (eds.), *NBER Macroeconomics Annual 2002,* (Cambridge, MA: MIT Press, 2003).

Svensson, Lars and Michael Woodford. "Indicator Variables for Optimal Policy," *Journal of Monetary Economics* 50 (2003), 691–720.

Svensson, Lars and Noah Williams. "Monetary Policy with Model Uncertainty: Distribution Forecast Targeting," unpublished manuscript, Princeton University (2005).

Swanson, Eric. "Optimal Nonlinear Policy: Signal Extraction with a Non-Normal Prior," *Journal of Economic Dynamics and Control* 30 (2006), 185–203.

Swanson, Eric. "Signal Extraction and Non-Certainty-Equivalence in Optimal Monetary Policy Rules," *Macroeconomic Dynamics* 8 (2004), 27–50.

Swanson, Eric, Gary Anderson, and Andrew Levin. "Higher-Order Perturbation Solutions to Dynamic, Discrete-Time Rational Expectations Models," *Federal Reserve Bank of San Francisco Working Paper* 2006–01 (2006).

Wieland, Volker. "Learning by Doing and the Value of Optimal Experimentation," *Journal of Economic Dynamics and Control* 24 (2000a), 501–34.

Wieland, Volker. "Monetary Policy, Parameter Uncertainty, and Optimal Learning," *Joural of Monetary Economics* 46 (2000b), 199–228.

Woodford, Michael. *Interest and Prices* (Princeton: Princeton University Press, 2003).

Zampolli, Fabrizio. "Optimal Monetary Policy in a Regime-Switching Economy," unpublished manuscript, Bank of England (2004).