Regime Switching, Learning, and the Great Moderation[*]

James Murray
Department of Economics
Indiana University[†]

April 4, 2007

Abstract

This paper examines empirically how the "bad luck" explanation can explain changing volatility in U.S. inflation and output when agents do not have rational expectations, but instead form expectations with least squares learning with an endogenously changing learning gain. Bad luck is modeled into a standard New Keynesian model by augmenting it with two states that evolve according to a Markov chain, where one state is characterized by large variances for structural shocks, and the other state has relatively smaller variances. Agents in a model are completely unaware of the state changing process and the underlying parameters of the model and instead form expectations by computing least squares regressions. Agents do suspect, however, that structural changes may occur, and so endogenously decide how much weight to give data further in the past based on the size of recent forecast errors relative to the average size of forecast errors. This type of endogenously changing learning mechanism can create periods of excess volatility without the need for changes in the variance of the underlying shocks. Therefore, when taking this model to the data, the degree of bad luck needed to explain volatile periods in U.S. history may not be as great as when assuming rational expectations or learning with a constant gain.

Keywords: Learning, regime switching, great moderation, New Keynesian model, maximum likelihood.
JEL classification: C13, E22, E31, E50.

---

1 Introduction

Figure 1 shows plots of the growth rate of real GDP and the GDP deflator annualized inflation rate for the United States since 1957. Aside from business cycle fluctuations, the data is characterized by changing periods of volatility. Macroeconomics has had difficulty explaining what has become to be called the Great Moderation, the period of excessive volatile during the late 1970s and early 1980s, followed by a relatively moderate period since the mid-1980s. This paper is also concerned with explaining the relatively moderate period during the 1960s, despite receiving less attention in the literature.

Leading explanations for these changes in volatility can be described as, good vs. bad policy, and good vs. bad luck. Lubik and Schorfheide (2004) find empirical evidence of a change in monetary policy from bad to good, occurring sometime between 1979 and 1982. They find that prior to 1979, monetary policy did not adjust the federal funds rate by more than one-to-one with inflation, and therefore under rational expectations, the equilibrium was indeterminate and the economy was subject to sunspot shocks which would lead to greater volatility. Milani (2005) estimates a similar model, but instead assumes that agents are not rational, but form expectations according to least squares learning. When splitting the sample at the same dates as Lubik and Schorfheide (2004), he finds little evidence of a change in monetary policy, finding the federal funds rate did adjust by more than one-to-one with inflation throughout the sample.

Primiceri (2005) incorporates learning and finds evidence of bad vs. good policy in an empirically founded model of unemployment rate and inflation determination. The "bad" policy in this work is actually characterized by optimal behavior by the central bank during the 1970s, but mis-perceptions by the central bank of the natural rate of unemployment and the degree of persistence of inflation lead to the wrong policy prescription. As time progresses and the central bank collects more data, it learns its mistake and eventually provides better policy, leading to period of great moderation.

Sims and Zha (2006) examine whether changes in U.S. volatility came from changes in luck or changes in policy by using a structural VAR with regime switching that only makes restrictions for identifying monetary policy, but otherwise leaves the model free of restrictions to allow for any interpretation of consumer and producer behavior. They find that the best fitting model is one with only bad luck. That is, the model that allows for no regime change

in policy or behavior, and only regime change in the variances of structural shocks.

In this paper, I examine a version of the "bad luck" story in the framework of a New Keynesian model where agents do not capable of rational expectations, but instead form expectations by running least squares regressions. I assume a simple model with only two states, one characterized by large variances for the structural shocks, and one characterized by relatively smaller variances. I assume the state evolves according to Markov chain were the current state is only dependent on the previous state, and the probabilities of switching between states is exogenous.

I assume agents do not have rational expectations, and are therefore completely unaware that there are two states of the economy, and that these states follow a Markov chain switching process. Agents do, however, remain suspect that structural changes may occur. They do have any knowledge about what types of structural changes are possible, or any idea with what probabilities structural changes can occur. I assume that expectations evolve according to the Marcet and Nicolini (2003) framework where the weight agents give observations that occurred further into the past is endogenous. Specifically, agents begin using a decreasing learning gain which consistent with ordinary least squares and no suspicion of structural change. However, if recent forecast errors become larger than the historical average, agents then suspect a structural change may have occurred and therefore increase their learning gain, giving larger weight to current observations which are believed to occur after a structural shock.

While a constant learning gain is theoretically capable of producing time varying volatility, in Murray (2007) I show in a New Keynesian model with no regime changes that constant gain learning is not able to deliver dynamics of U.S. inflation and output much better than rational expectations, and both frameworks especially fail to explain the excess volatility of the 1970s. Milani (2007) simulates a model estimated with U.S. data that includes the endogenously changing learning gain suggested by Marcet and Nicolini (2003) and finds that this type of learning can effectively explain significant changes in volatility. Moreover, his estimates imply that during most of the 1970s decade, agents suspected structural change and therefore used a high constant gain.

This paper builds on Milani's analysis by also allowing for recurrent regime changes in the volatility of the structural shock. Estimation of this model will decompose of the explanation for changes in volatility into changes in the volatility of structural shocks, and endogenous

changing in the learning gain. Since both of these mechanisms deliver greater volatility, I expect the results to indicate that periods of excessive volatility are explained by a mixture of changes in learning and changes in luck. Furthermore, I expect that when failing to allow for changes in the learning gain one will over estimate both the frequency of the "bad" regime occurring, and the volatility of the shocks that occur during the bad regime.

2 Model

The crucial extentions of an endogenously changing learning gain and a Markov chain process for changes in structural volatility are incorporated into a standard New Keynesian model of output and inflation determination and monetary policy. There are a continuum of consumers types on the unit interval, and a continuum of intermediate goods producers on the same unit interval. Each consumer type as a specific labor skill that is only hired by the corresponding intermediate goods producer. It is assumed there are a number of consumers of each type, so that no consumer has market power over the wage, and that there are an equal number of consumers in each type, so that relative output levels of intermediate goods do not depend on the distribution of consumer types.

There is one final good that is used for consumption, and produced using all the intermediate goods. The intermediate goods are imperfect substitutes for eachother in production, therefore intermate goods firms are monopolistically competitive. Prices for intermediate goods are subject to a Calvo (1983) pricing friction where only a fraction of firms are able to re-optimize their price every period, and the firms fortunate to do so is randomly determined, independently of firms' histories.

2.1 Consumers

Each consumer type has a specific labor skill that can only be hired by a specific intermediate goods producing firm. Since each intermediate goods firm has a different labor demand, wage income will be different for each consumer type. Given a perfect asset market, though, consumption will be equal across all consumers. Each consumer type $i \in (0, 1)$ maximizes utility,

$$E_0^* \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1 - \frac{1}{\sigma}} \xi_t \left(c_t - \eta c_{t-1} \right)^{1 - \frac{1}{\sigma}} - \frac{1}{1 + \frac{1}{\mu}} n_t(i)^{1 + \frac{1}{\mu}} \right],$$

subject to the budget constraint,

$$c_t + b_t(i) = \frac{1 + r_{t-1}}{1 + \pi_t} b_{t-1}(i) + \frac{w_t(i)}{p_t} n_t(i) + \Pi_t - \tau_t$$

where the expectation operator E_t^* denotes possibly non-rational expectations. Consumption at time t, given by c_t, is not indexed by individual type i since it is equal across all agents. The remaining variables include ξ_t which is an aggregate preference shock, $n_t(i)$ and $w_t(i)/p_t$ are the labor supply and real wage of individual i at time t, respectively, $b_t(i)$ is individual i's purchase of real government bonds at time t, r_t is the nominal interest rate paid on government bonds, π_t is the inflation rate of the price of the final good, Π_t is the value of profits earned by owning stock in firms, and τ_t is the value of real lump sum taxes. The preference parameters are $\sigma \in (0, \infty)$, which is the intertemporal elasticity of substitution, $\mu \in (0, \infty)$ which is the elasticity of labor supply, and $\eta \in [0, 1)$, which is the degree of habit formation.

Habit formation is added to the model because it introduces a source of consumption (and therefore output) persistence that has been found to be significant in rational expectations models. For example Smets and Wouters (2005) estimate a rational expectations New Keynesian model with numerous for both the United States and Euro area and find point estimates for the degree of habit formation very close to 1. Furthermore, Fuhrer (2000) finds that habit formation leads to "hump shaped" impulse response functions that can be supported by the data. The importance of habit formation may not be so important when dropping the assumption of rational expectations. Milani (2005) estimates a New Keynesian model with learning, and shows that under learning, the degree of habit formation falls close to zero.

Log-linearizing the consumers' first order conditions leads to the following Euler equation,

$$\hat{\lambda}_t = E_t^* \hat{\lambda}_{t+1} + \hat{r}_t - E_t^* \pi_{t+1}, \tag{1}$$

where a hat indicates the percentage deviation of the variable from its steady state.[1] Here,

---

[1] A hat is omitted from inflation because it will be necessary to assume the steady state inflation rate is equal to zero when log-linearizing the firms' profit maximizing conditions.

$\hat{\lambda}_t$ is the marginal utility of real income, given by,

$$\hat{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)} \left[\beta\eta\sigma E_t^* \hat{c}_{t+1} - \sigma(1+\beta\eta^2)\hat{c}_t + \sigma\eta\hat{c}_{t-1} \right] + \left(\hat{\xi}_t - \beta\eta E_t^* \hat{\xi}_{t+1} \right). \quad (2)$$

2.2 Production

There is one final good used for consumption and investment which is sold in a perfectly competitive market and produced with a continuum of intermediate goods. The production function is given by,

$$y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}} \quad (3)$$

where y_t is the output of the final good, $y_t(i)$ is intermediate good i, and $\theta \in (1, \infty)$ is the elasticity of substitution in production. Profit maximization leads to the demand for each intermediate good,

$$y_t(i) = \left[\frac{p_t(i)}{p_t} \right]^{-\theta} y_t, \quad (4)$$

where $p_t(i)$ is the price of intermediate good i and p_t is the price of the final good. Substituting equation (4) into (3) leads to a consumption price index that holds in equilibrium,

$$p_t = \left[\int_0^1 p_t(i)^{1-\theta} di \right]^{\frac{1}{1-\theta}}. \quad (5)$$

Each intermediate good is produced according the constant returns to scale production function $y_t(i) = z_t n_t(i)$, where z_t is an exogenous technology shock common to all firms. Given a level of production $y_t(i)$, firms choose labor demand to minimize total cost $\frac{w_t(i)}{p_t} n_t(i)$. When labor markets clear, it can be shown that firm i's optimal choice for labor leads to the log-linearized marginal cost of firm i equal to,

$$\hat{\psi}_t(i) = \frac{1}{\mu}\hat{y}_t(i) - \hat{\lambda}_t - \left(\frac{1}{\mu} + 1 \right) \hat{z}_t. \quad (6)$$

Summing equation (6) across all firms leads to the average marginal cost in the economy,

$$\hat{\psi}_t = \frac{1}{\mu}\hat{y}_t - \hat{\lambda}_t - \left(\frac{1}{\mu} + 1 \right) \hat{z}_t. \quad (7)$$

Firms' pricing decisions are subject to the Calvo (1983) pricing friction, where only a

constant fraction of firms are able to re-optimize their in a given period. I suppose that firms who are not able to re-optimize their price may adjust their price by a fraction, γ, of the previous period's inflation rate. Let $\omega \in (0,1)$ denote the fraction of firms who are not able to re-optimize their prices each period. Since these specific firms are randomly determined, ω^T is the probability a firm will not be able to re-optimize its price for T consecutive periods. A firm who is able to re-optimize its price maximizes the following present discounted utility value of profits earned while the firm is unable to re-optimize its price again:

$$E_t^* \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_t(i)\pi_{t+T}^*}{p_{t+T}} \right) y_{t+T}(i) - \Psi\left[y_{t+T}(i) \right] \right\},$$ (8)

where $\Psi\left[y_{t+T}(i) \right]$ is the real total cost function of producing $y_{t+T}(i)$ units, given the optimal decision for labor, and $\pi_{t+T}^* = \prod_{j=1}^{T}(1+\gamma\pi_{t+j-1})$ is degree to which the firm's price is able to be adjusted according to inflation indexation. It can be shown that the first order condition for $p_t(i)$ combined with the final good price index, equation (5), leads to the log-linear Phillips equation,

$$\pi_t = \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{\beta}{1+\beta\gamma}E_t^*\pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega(1+\theta\mu)(1+\beta\gamma)}\hat{\psi}_t$$ (9)

2.3   Monetary Policy

The nominal interest rate is determined jointly with output and inflation by monetary policy. In this paper I assume the monetary authority follows a Taylor (1993) type rule of the form,

$$\hat{r}_t = \rho_r\hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi\pi_t + \psi_y\hat{y}_t \right) + \epsilon_{r,t}$$ (10)

where $\rho_r \in [0,1)$ is a degree of interest rate smoothing desired by the monetary authority, $\psi_\pi \in (0,\infty)$ is the feedback on the interest rate to inflation, $\psi_y \in (0,\infty)$ is the feedback on the interest rate to deviations of output from the steady state, and $\epsilon_{r,t}$ is an independently and identically distributed exogenous monetary policy shock.

2.4   Regime Switching

The bad luck explanation for why the United States experienced periods of excessive volatility is that the variances of the structural shocks were larger during these periods. To model

this explanation in the framework of the New Keynesian model I suppose the variance of the technology shock, z_t, and the preference shock, ξ_t, are determined by two states. Let state 1, denoted by s_1, be the state where the shocks are relatively less volatile, and state 2, denoted by s_2, be the state where the shocks are relatively more volatile. I furthermore assume that these shocks may be serially correlated, so they evolve according to,

$$z_t = \rho_z z_{t-1} + \epsilon_{z,t}(s_t) \tag{11}$$

$$\xi_t = \rho_\xi \xi_{t-1} + \epsilon_{\xi,t}(s_t) \tag{12}$$

where the innovation of the shocks depend on the state. The innovations in a given state will be independently normally distributed with mean zero and variances given by,

$$Var\left(\begin{bmatrix} \epsilon_{z,t}(s_t) \\ \epsilon_{\xi,t}(s_t) \end{bmatrix}\right) = \begin{cases} \begin{bmatrix} \sigma_{z,1}^2 & 0 \\ 0 & \sigma_{\xi,1}^2 \end{bmatrix}, & \text{if } s_t = 1 \\ \begin{bmatrix} \sigma_{z,2}^2 & 0 \\ 0 & \sigma_{\xi,2}^2 \end{bmatrix}, & \text{if } s_t = 2 \end{cases} \tag{13}$$

The state s_t evolves according to a two state Markov chain. Let $p_j \in (0,1)$ denote the probability of staying in state j at time t, given the economy is at state j at time $t-1$, for $j \in \{1,2\}$. This implies that the probability of moving from state j in $t-1$ to the other state in time t is given by $1 - p_j$. The state then evolves according to the transition matrix,

$$P = \begin{bmatrix} p_1 & 1 - p_1 \\ 1 - p_2 & p_2 \end{bmatrix}. \tag{14}$$

3   Learning

I deviate from the rational expectations assumption and assume agents form expectations by computing least squares regressions and using these forecasts as their expectations. I also assume that agents have no knowledge of the regime switching process and no knowledge about what types of regimes are possible or with what probabilities. I do allow agents to adjust their expectations if they suspect structural changes have occurred. One way to model this is to use constant gain learning, which is consistent with agents' forecasts

based on weighted least squares, where the weights decline geometrically with the age of the observations. With a constant gain, the weight put on the latest observation is always the same, regardless of how much data the agents already have for their forecasts. One benefit of this type of learning is that learning dynamics persist in the long run. If agents instead use ordinary least squares (OLS), then the weight agents give to additional observations diminishes as their sample size approaches infinity. Said another way, the learning gain decreases and approaches zero as time approaches infinity. Learning dynamics therefore disappear completely in the long run when expectations are formed with OLS.

Authors such as Sargent (1999) and Evans and Honkapohja (2001) have suggested constant gain learning is a natural way to model agents which have a constant suspicion that their may have been a structural change. Therefore every period agents give the most weight to the latest observations, and give lesser weight the longer ago an observation occurred. Constant gain learning can in theory lead to time varying volatility of expectations, and therefore time varying volatility of outcomes for inflation and output. This affect depends on the size of the constant gain, and for empirically plausible values, this type of learning fails to deliver very big effects. Williams (2003) finds this to be the case in a simple model with simulated data. In Murray (2007), I show in an esimated New Keynesian model that constant gain learning provides dynamics very similar to rational expectations.

Marcet and Nicolini (2003) suggest an alternative way to model dynamic expectations. Instead of assuming that agents never suspect a structural change, as is consistent with OLS, or assuming that agents always suspect a structural change with equal likelihood, as is consistent with a constant gain, they take a mixture of these methods where the learning gain changes endogenously. Suppose agents begin with no suspicion of structural change. As time progresses they form their expectations using a decreasing learning gain consistent with OLS. Agents then begin to suspect a structural change if recent forecast errors are larger than historical averages. When this happens, agents switch to a relatively large learning gain. As long as forecast errors are large, the learning gain remains at this high point. When forecast errors start becoming small, the learning gain then starts decreasing from this high point at a rate consistent with OLS. I explain explicitly this endogenous gain algorithm below, but first it is first necessary to explain, specific details of least squares learning in the framework of a dynamic stochastic general equilibrium model like the New Keynesian model in Section 2.

3.1   Least Squares Learning

The log-linearized model in Section 2 can be expressed in the general form:

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi \epsilon_t(s_t), \tag{15}$$

where x_t is a vector of minimum state variables (expressed in terms of percentage deviations form the steady state), all observable to agents in the model, and ϵ_t is a vector of structural shocks, which as discussed in the previous section, may be state dependent. For the New Keynesian model examined in this paper $x_t = [\hat{y}_t \ \pi_t \ \hat{r}_t \ \hat{z}_t \ \hat{\xi}_t]'$, and $\epsilon_t(s_t) = [\epsilon_{z,t}(s_t) \ \epsilon_{\xi,t}(s_t) \ \epsilon_{r,t}]'$. The minimum state variable solution of the model implies the rational expectation for x_{t+1} is given by,

$$E_t x_{t+1} = G x_t, \tag{16}$$

where the elements of the matrix G are a function of the parameters of the model and may be determined by the method of undetermined coefficients.[2] Agents that form expectations by learning do not know any of the structural parameters and instead estimate the matrix G by least squares.

The timing in a given period of how agents make expectations and decisions is as follows. At the begining of period t agents wake up with data realized through period $t-1$. They collect this data and use their least squares results to form an estimate for G which they then use to make forecasts for future realizations of variables such as output and inflation. Given these expectations, agents consumption and pricing decisions are implemented, and only then do time t outcomes become realized. In the next period, these outcomes are then available as data to the agents, and the process begins again.

There is no constant term in the general form of the model, equation (15), or in the rational expectation, given in equation (16), since all variables are expressed in terms of percentage deviations from the steady state. Since agents are not endowed with information about the parameters of the model, it is realistic to suppose that agents do not know steady

---

[2]Log-linearizing the model eliminates any importance agents place on risk that is inherent in a non-linear solution of the model. Equation (16) is determined in a linear model and so is completely independent of the variance of the structural shocks. In the non-linear model, risk may play an important role in agents behavior especially in the framework of regime switching volatility, but I ignore this avenue and suppose that changes in agents behavior not explained by the first order properties come from changes in expectations due to learning.

state values. Therefore, I suppose that agents also estimate a constant term in equation (16). Let $\hat{G}_t^*$ denote agents' time t estimate for matrix G which also includes a column for a constant term so that. Therefore, $\hat{G}_t^* = [\hat{g}_{0,t} \ \hat{G}_t]$, where $\hat{g}_{0,t}$ is the time t estimate of the constant term and $\hat{G}_t$ is the time t estimate for G.

If agents use OLS, then,

$$\left(\hat{G}_t^*\right)' = \left(\frac{1}{t-1}\sum_{\tau=1}^{t-1} x_{\tau-1}^* x_{\tau-1}^{*'}\right)^{-1}\left(\frac{1}{t-1}\sum_{\tau=1}^{t-1} x_{\tau-1}^* x_{\tau}'\right), \tag{17}$$

where $x_t^* = [1 \ x_t']'$ is the vector of regressors agents use, which is simply x_t with a 1 augmented in order to estimate the constant term. Agents then form the expectation,

$$E_t^* x_{t+1} = \hat{g}_{0,t} + \hat{G}_t E_t^* x_t = (I + \hat{G}_t)\hat{g}_{0,t} + \hat{G}_t^2 x_{t-1}. \tag{18}$$

The least squares estimate for G_t^* can be rewritten as the following recursive form:

$$\hat{G}_t^* = \hat{G}_{t-1}^* + g_t(x_{t-1} - \hat{G}_{t-1}^* x_{t-2}^*)x_{t-2}^{*'} R_t^{-1}, \tag{19}$$

$$R_t = R_{t-1} + g_t(x_{t-2}^* x_{t-2}^{*'} - R_{t-1}), \tag{20}$$

where $g_t = 1/(t-1)$ is the learning gain.[3] The recursive form shows precisely how expectations are adaptive. The term enclosed in parentheses in equation (19) is the realized forecast error for the previous estimate $\hat{G}_{t-1}^*$. The degree to which agents adjust their increases with the size of this forecast error, depending on the variance of the estimated coefficients, captured by the matrix R_t, and depending on the size of the learning gain, g_t. The larger is the learning gain, the more expectations respond to the latest forecast error. When agents use OLS, g_t approaches zero as time approaches infinity. Under constant gain learning, g_t remains at some constant level g, so that the degree to which new observations can affect expectations is always the same.

Under the mixed learning framework suggested by Marcet and Nicolini (2003), the learning gain g_t decreases over time, but may endogenously jump to a higher level when forecast errors are especially large. Let $\alpha_t \equiv 1/g_t$ be the inverse of the learning gain, which under ordinary least squares is interpreted as the sample size. The learning gain is assumed to

---

[3]To show this, let $R_t = \frac{1}{t-1}\sum_{\tau=1}^{t-1} x_{\tau-1}^* x_{\tau-1}^{*'}$ and $\left(\hat{G}_t^*\right)' = R_t^{-1}\left(\frac{1}{t-1}\sum_{\tau=1}^{t-1} x_{\tau-2}^* x_{\tau-1}'\right)$.

evolve according to,

$$
\alpha_t = \begin{cases} \alpha_{t-1} + 1 & \text{if } \dfrac{1}{J} \sum_{j=1}^{J} \sum_{v=1}^{n} \left| \dfrac{x_{t-j}(v) - \hat{G}^*_{t-j}(v) x^*_{t-j-1}}{\hat{G}^*_{t-j}(v) x^*_{t-j-1}} \right| < \nu \\ \\ \alpha & \text{otherwise} \end{cases} \tag{21}
$$

where n denotes the number of variables in the model, $x_{t-j}(v)$ denotes the vth element of x_{t-j} and $\hat{G}^*_{t-j}(v)$ denotes the vth row of $\hat{G}^*_{t-j}(v)$, which is used to forecast variable v. The parameter J is the number of recent periods agents look at when deciding to change their learning gain. Marcet and Nicolini (2003) assume $J = 1$, so that agents are concerned with only the most recent forecast error. I will fix $J = 4$, so for quarterly data agents examine the forecast errors for the last year and adjust the learning gain if the forecast error is too large. Finally, the parameter $\nu \in (0, \infty)$ is the threshold for how large the forecast errors must be to induce agents to increase their learning gain. Notice that the term in absolute value in equation 21 is the forecast error for variable v as fraction of the forecast for variable v. Since forecast errors for each variable v is given by a percentage, they can be added up over all the variables the agents make forecasts for. This learning mechansim nests the special cases when agents always use OLS or always use a constant gain. To restrict agents to always use OLS, the parameter ν can be set to zero. To restrict agents to always use a constant gain, the parameter ν is set to infinity.

The learning parameters that govern the evolution of expectatations are therefore α and ν. In the estimation procedure in below, I estimate ν and the constant learing gain $g \equiv 1/\alpha$.

4   Estimation

Substituting the least squares expectation of time $t + 1$ variables, given in equation (18), into the general form of the dynamic stochastic general equilibrium model, given in equation (15), leads to the following regime dependent state equation,

$$
x_t = b_t + F_t x_{t-1} + v_t(s_t) \tag{22}
$$

where,

$$
b_t = \Omega_0^{-1} \Omega_2 \left(I + \hat{G}_t \right) \hat{g}_{0,t} \tag{23}
$$

$$F_t = \Omega_0^{-1}\left(\Omega_1 + \Omega_2 \hat{G}_t^2\right), \tag{24}$$

$$v_t(s_t) = \Omega_0^{-1}\Psi\epsilon_t(s_t). \tag{25}$$

The variance of $v_t(s_t)$ is regime dependent and given by,

$$Var\left[v_t(s_t)\right] \equiv Q(s_t) = \left(\Omega_0^{-1}\Psi\right)Var\left[\epsilon_t(s_t)\right]\left(\Omega_0^{-1}\Psi\right)', \tag{26}$$

where $Var\left[\epsilon_t(s_t)\right]$ is given in equation (13). The matrix $\hat{G}_t$ and vector $\hat{g}_{0,t}$ are determined by the learning algorithm described in equations (19, (20), and (21). The evolution of s_t is determined by the transition matrix, given in equation (14). The state equation is mapped to the data with the following system of observation equations,

$$\begin{aligned} GDP_t &= 100\hat{y}_t, \\ INF_t &= \pi^* + 400\pi_t, \\ FF_t &= r^* + 400\hat{r}_t, \end{aligned} \tag{27}$$

which can be written in the compact form,

$$x_t^{obs} = a + H'x_t, \tag{28}$$

where a and H are the associated vector and matrix that correspond to equation (27).

Equations (22) and (27) are in a standard state-space representation. The Kalman fiter procedure described by Hamilton (1994) combined with Hamilton's 1989 procedure for handling regime switching can be used to map the regime dependent state equation and observation equation into a likelihood function. I'm still working out these details.

TO BE CONTINUED...

References

CALVO, G. A. (1983): "Staggered prices in a utility maximizing framework," Journal of Monetary Economics, 12, 383–398.

EVANS, G. W., AND S. HONKAPOHJA (2001): Learning and Expectations in Macroeconomics. Princeton University Press.

FUHRER, J. C. (2000): "Habit formation in consumption and its implications for monetary-policy models," American Economic Review, 90, 367–390.

HAMILTON, J. (1989): "A new approach to the economic analysis of nonstationary time series and the business cycle," Econometrica, 57, 357–384.

——— (1994): Time Series Analysis. Princeton University Press.

LUBIK, T., AND F. SCHORFHEIDE (2004): "Testing for indeterminacy: An application to U.S. monetary policy," American Economic Review, 94, 190–217.

MARCET, A., AND J. P. NICOLINI (2003): "Recurrent Hyperinflations and learning," American Economic Review, 93, 1476–1497.

MILANI, F. (2005): "Expectations, learning and macroeconomic persistence," Working Paper.

——— (2007): "Learning and time-varying macroeconomic volatility," Working Paper.

MURRAY, J. M. (2007): "Empirical significance of learning in a New Keynesian model with firm-specific capital," Working Paper.

PRIMICERI, G. E. (2005): "Why inflation rose and fell: policymakers' beliefs and US postwar stabilization policy," NBER Working Paper 11147.

SARGENT, T. J. (1999): Conquest of American Inflation. Princeton University Press.

SIMS, C. A., AND T. ZHA (2006): "Were there regime switches in U.S. monetary policy," American Economic Review, 96, 54–81.

SMETS, F., AND R. WOUTERS (2005): "Comparing shocks and friction in U.S. and Euro area business cycles: A Bayesian DSGE approach," Journal of Applied Econometrics, 20, 161–183.

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

WILLIAMS, N. (2003): "Adaptive learning and business cycles," Working Paper.

Figure 1: U.S. Output Growth and Inflation 1957:Q1 - 2005:Q4

Output Growth

Inflation