CAEPR Working Paper
#2008-017

Initial Expectations in New Keynesian Models with Learning

James Murray
Indiana University Bloomington

June 13, 2008

Initial Expectations in New Keynesian Models with Learning*

James Murray
Department of Economics
Indiana University†

June 8, 2008

Abstract

This paper examines how the estimation results for a standard New Keynesian model with constant gain least squares learning is sensitive to the stance taken on agents beliefs at the beginning of the sample. The New Keynesian model is estimated under rational expectations and under learning with three different frameworks for how expectations are set at the beginning of the sample. The results show that initial beliefs can have an impact on the predictions of an estimated model; in fact previous literature has exposed this sensitivity to explain the changing volatilities of output and inflation in the post-war United States. The results indicate statistical evidence for adaptive learning, however the rational expectations framework performs at least as well as the learning frameworks, if not better, in in-sample and out-of-sample forecast error criteria. Moreover, learning is not found to better explain time varying macroeconomic volatility any better than rational expectations. Finally, impulse response functions from the estimated models show that the dynamics following a structural shock can depend crucially on how expectations are initialized and what information agents are assumed to have.

Keywords: Learning, expectations, New Keynesian model, maximum likelihood.
JEL classification: C13, E31, E50.

*I am grateful for the advice and guidance of Eric Leeper, Kim Huynh, Brian Peterson, and Todd Walker; for useful conversations with Fabio Milani and Bruce Preston; and for comments by the participants of Indiana University economics department seminars. All errors are my own.

†*Mailing address*: 100 S Woodlawn, Bloomington, IN 47405. *E-mail address*: jmmurray@indiana.edu. *Phone number*: (574)315-0459.

1 Introduction

Recently there has been a growing amount of literature concerning the effects of least squares learning, a type of adaptive expectations mechanism, on empirical puzzles encountered in monetary economics. Least squares learning is an expectations framework where agents in a model do not know the parameters that govern the economy and therefore form expectations by collecting past data and computing forecasts from least squares estimation results. Orphanides and Williams (2005b) show with a simple calibrated model and simulated impulse response functions that such a learning framework can cause prolonged periods of inflation, that would not occur under rational expectations, following an inflation shock. Learning has also been suggested to be responsible for the slowdown in macroeconomic volatility since the middle 1980s, a phenomenon commonly referred to as the Great Moderation. For example, Orphanides and Williams (2005a) suggest in another paper that the monetary authority forms their expectations by learning and was under-estimating the natural rate of unemployment during the 1970s, causing an incorrect prescription for expansionary monetary policy. Primiceri (2006) takes this argument further and suggests that over the course of the 1970s and early 1980s the monetary authority gradually gained precision in their estimates, causing policy prescription to correctly adjust to stabilize output and inflation.

Milani (2007) has suggested that learning can better explain persistence in output and inflation in the context of a New Keynesian model better than traditional means of modeling persistence such as habit formation and inflation indexation. Milani also estimates the size of the constant learning gain, the parameter responsible for the degree to which expectations evolve, and finds that expectations are adaptive over a post-war sample period.

The results in Milani (2007) and Primiceri (2006) depend on calibrated values for expectations at the beginning of the sample. Moreover, the dynamics predicted by learning can be influenced by the assumptions regarding agents information sets. The purpose of this paper is to examine the role constant gain learning, a specific type of least squares learning, has on the predictions of an estimated standard New Keynesian model. Moreover, this paper carefully considers different frameworks for how initial expectations are specified and what information agents are able to collect to form their forecasts. Through examining forecast er-

rors, rational expectations is shown to explain the data nearly as well, if not better, than the various learning frameworks. Even so, the estimates for the learning gain indicate statistical evidence for adaptive expectations. Impulse response functions are examined to determine the effects the various learning frameworks have on the dynamics of the model following a structural shock. The results indicate that the impulse response functions can vary depending on the assumptions for agents' information sets and initial expectations. Moreover, the findings indicate that learning can lead to some prolonged effects in output and inflation following a structural shock.

The results do not confirm, however, previous literature that suggests learning can explain periods of excessive volatility in inflation and output followed by the subsequent decline in volatility. Evolution of the forecast errors over the sample indicate the rational expectations model and learning models all make similar errors, and all models make the largest errors during the 1970s and early 1980s when inflation and output were especially volatile.

The next section describes the basic setup of the New Keynesian model. Section 3 describes the learning procedure and how learning is incorporated into a standard linear dynamic stochastic general equilibrium (DGSE) model. Section 4 describes the estimation procedure and the issues involved in initializing expectations and determining agents' information sets. Section 5 presents the results, and Section 6 concludes.

2 Model

Learning is examined within the context of a standard New Keynesian model. The New Keynesian model is one of the most commonly used models in monetary economics as it provides a convenient framework to examine theoretical and empirical issues for monetary policy and inflation and output determination. This section describes the set-up and log-linearization of the rational expectations model and in the next section the rational expectations are replaced by expectations under learning.[1]

[1]This is perhaps the most common way to incorporate learning into dynamic macroeconomic models. However, as Marcet and Sargent (1989) point out and Preston (2005) further demonstrates, this method is not consistent with learning in the microfoundations of the model because the least squares expectations operator does not follow the law of iterated expectations, a property that is assumed when solving for the log-linear solution.

The model consists of three sectors that describe consumer behavior, producer behavior under imperfectly flexible prices, and monetary policy. The first sector is an equation or system of equations that describes optimal consumer behavior. When this sector can be conveniently written in one equation, this is often called the "IS equation". The second sector is a single equation, referred to as the Phillips curve, that describes optimal producer behavior when firms are subject to a pricing friction. The final sector is the monetary authority, which is usually assumed to follow a simple nominal interest rate rule. The sectors jointly determine the dynamics of the output gap (the percentage difference between real GDP and potential GDP), the inflation rate, and the nominal interest rate.

2.1 Consumers

There are a continuum of consumer types and a continuum of intermediate good producers, each on the unit interval. Each consumer type has a specific type of labor skill that can only be hired by a corresponding intermediate good firm. It is assumed that there many consumers of each type so that no consumer has market power over their wage. Moreover, it is assumed that there are the same number of consumers in each type, so that the output levels of intermediate goods do not depend on the distribution of consumer types. Different intermediate goods firms may pay different wages, so labor income may be different for each consumer type. To simplify the model, it is further assumed that there is a perfect asset market so despite differences in labor income, all consumers choose the same level of consumption.

Each consumer of type $i \in (0,1)$ chooses consumption, c_t, labor supply, $n_t(i)$, and purchases of real government bonds, $b_t(i)$, to maximize lifetime utility,

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

subject to the budget constraint,

$$c_t + b_t(i) = \frac{1+r_{t-1}}{1+\pi_t} b_{t-1}(i) + \frac{w_t(i)}{p_t} n_t(i) + \Pi_t - \tau_t. \tag{2}$$

where ξ_t is an aggregate preference shock, $w_t(i)/p_t$ is the real wage paid to type i labor; Π_t is the total value of profits consumers earn by owning stock in firms, and τ_t is the real value of lump sum taxes. The preference parameters are the intertemporal elasticity of substitution, denoted by $\sigma \in (0, \infty)$; the elasticity of labor supply, denoted by $\mu \in (0, \infty)$; and the degree of habit formation, denoted by $\eta \in [0, 1)$.

When the degree of habit formation is greater than zero, consumers' utility from current consumption depends on their previous level of consumption. Habit formation introduces persistence in consumption, and therefore output. Significant output persistence is commonly found in empirical studies of DSGE models. For example, Smets and Wouters (2005) find point estimates of habit formation close to unity. Furthermore, Fuhrer (2000) finds that habit formation leads to "hump-shaped" impulse response functions, a characteristic commonly supported by U.S. and European data. Milani (2007) finds a significant degree of habit formation, but only under rational expectations. When estimating the model with constant gain learning, he finds an estimate for the degree of habit formation close to zero.

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

$$\hat{\lambda}_t = E_t \hat{\lambda}_{t+1} + \hat{r}_t - E_t \pi_{t+1}, \tag{3}$$

where $\hat{\lambda}_t$ is the percentage deviation from the steady state of the Lagrange multiplier on the budget constraint, (2), and is therefore interpreted as the marginal utility of real income. A hat indicates the percentage deviation of a variable from its steady state.[2] Utility maximization leads to the following log-linear marginal utility of income,

$$\hat{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)} \left[\beta\eta\sigma E_t \hat{c}_{t+1} - \sigma(1+\beta\eta^2)\hat{c}_t + \sigma\eta\hat{c}_{t-1}\right] + \left(\hat{\xi}_t - \beta\eta E_t \hat{\xi}_{t+1}\right). \tag{4}$$

The marginal utility of income, (4), and the Euler equation, (3), make up the IS sector of the model.

[2] A hat is omitted from π_t because it is necessary to assume the steady state level of inflation is equal to zero when deriving the log-linear supply relationship.

2.2 Producers

There is one final good used for consumption which is sold in a perfectly competitive market and produced with a continuum of intermediate goods according to the production function,

$$y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di\right]^{\frac{\theta}{\theta-1}}, \tag{5}$$

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

$$y_t(i) = \left[\frac{p_t(i)}{p_t}\right]^{-\theta} y_t, \tag{6}$$

where $p_t(i)$ is the price of intermediate good i and p_t is the price of the final good. Substituting equation (6) into equation (5) leads to the following expression for the price of the final good in terms of the prices of intermediate goods,

$$p_t = \left[\int_0^1 p_t(i)^{1-\theta} di\right]^{\frac{1}{1-\theta}}. \tag{7}$$

Each intermediate good is sold in a monopolistically competitive market and is produced according to the production function, $y_t(i) = z_t n_t(i)$, where z_t is an aggregate technology shock. It can be shown that intermediate goods firms' optimal choices for labor demand and labor market clearing leads to the following aggregate log-linear marginal cost,

$$\hat{\psi}_t = \frac{1}{\mu}\hat{y}_t - \hat{\lambda}_t - \left(\frac{1}{\mu} + 1\right)\hat{z}_t. \tag{8}$$

Firm's pricing conditions are subject to the Calvo (1983) pricing friction, where only a constant fraction of firms are able to re-optimize their price in a given period. The firms that are able to re-optimize their price is randomly determined, completely independently of firms' prices or any other characteristics or history. I suppose that firms who are not able to re-optimize their price do adjust their price by a fraction, $\gamma \in [0, 1)$, of the previous period's inflation rate. A positive degree of price indexation introduces a source of persistence in

inflation which is often found to be statistically significant when estimating New Keynesian models (see for example, Smets and Wouters (2003), (2003), (2007), and Milani (2007)).

Let $\omega \in (0,1)$ denote the fraction of firms that are not able to re-optimize their prices every period. Since these firms are randomly determined, ω^T is the probability that a firm will not be able to re-optimize its price for T consecutive periods. A firm who is able to re-optimize chooses its price to maximize the following present discounted utility value of profits earned while the firm is unable to re-optimize its price again:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_t(i)\pi^*_{t+T}}{p_{t+T}} \right) y_{t+T}(i) - \Psi\left[y_{t+T}(i)\right] \right\}, \tag{9}$$

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 adjust 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 (7), leads to the log-linear Phillips equation[3],

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{\mu(1-\omega)(1-\omega\beta)}{\omega(\mu+\theta)}\hat{\psi}_t\right]. \tag{10}$$

2.3 Fully Flexible Prices

The IS equations and Phillips equations can be re-written in terms of the difference from the outcome under fully flexible prices. This allows the model to be taken to data on the output gap, the percentage deviation of real GDP from real potential GDP, as measured by the Congressional Budget office.

Let $\tilde{y}_t = \hat{y}_t - \hat{y}^f_t$ and $\tilde{\lambda}_t = \hat{\lambda}_t - \hat{\lambda}^f_t$ denote the percentage deviation of output and marginal utility from their fully flexible price outcomes, where a superscript f denotes the outcome under fully flexible prices. Under flexible prices the linearized Euler equation, (3), and marginal utility of income, (4), still hold. Using these conditions and imposing goods

[3]It is assumed during the log-linearization that there is a steady state level of the price level, which implicitly assumes the steady state level of inflation is equal to zero.

market clearing that consumption is equal to output implies,

$$\tilde{\lambda}_t = E_t\tilde{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1} - r_t^n, \tag{11}$$

$$\tilde{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\tilde{y}_{t+1} - \sigma(1+\beta\eta^2)\tilde{y}_t + \sigma\eta\tilde{y}_{t-1}\right], \tag{12}$$

where r_t^n is the percentage deviation of the natural interest rate from its steady state. The "natural interest rate" is the interest rate that would occur under fully flexible prices. I suppose that r_t^n follows the stochastic exogenous process,

$$r_t^n = \rho_n r_{t-1}^n + \epsilon_{n,t}, \tag{13}$$

where $\epsilon_{n,t}$ is an independently and identically distributed shock.

When prices are fully flexible, it can be shown that intermediate goods firms will all choose the same price in a given period, and the marginal cost of production is constant, and therefore always will be equal to its steady state value. Under fully flexible prices, equation (8) implies,

$$\hat{\psi}_t^f = \frac{1}{\mu}\hat{y}_t^f - \hat{\lambda}_t^f - \left(\frac{1}{\mu}+1\right)\hat{z}_t = 0.$$

One can solve this equation for $\hat{z}_t$ and substitute it back into the equation for marginal cost, (8). Plugging this expression for marginal cost into equation (10) yields the following Phillips curve in terms of the output gap,

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega(\mu+\theta)}(\tilde{y}_t - \mu\tilde{\lambda}_t)\right].$$

While this expression for the Phillips curve is not subject to a structural shock, when estimating the model by maximum likelihood it is convenient to have a shock here to avoid the problem of stochastic singularity. The Phillips curve is amended with a "cost-push" shock so the form that is estimated is given by,

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \kappa(\tilde{y}_t - \mu\tilde{\lambda}_t) + u_t\right], \tag{14}$$

where κ is the reduced form coefficient on the marginal cost and u_t is an exogenous cost-push shock that evolves according to,

$$u_t = \rho_u u_{t-1} + \epsilon_{u,t}, \tag{15}$$

where $\epsilon_{u,t}$ is an independently and identically distributed shock.

2.4 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 where the interest rate is set in response to expected output and inflation, with a preference for interest rate smoothing, according to,

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi E_t \pi_{t+1} + \psi_y E_t \tilde{y}_{t+1}\right) + \epsilon_{r,t} \tag{16}$$

where $\rho_r \in [0,1)$ is the degree of exogenous interest rate persistence, $\psi_\pi \in (0,\infty)$ is the degree to which monetary policy responds to expectations of future inflation above the steady state level of inflation, $\psi_y \in (0,\infty)$ is the degree to which monetary policy responds to the expected output gap, and $\epsilon_{r,t}$ is an independently and identically distributed exogenous monetary policy shock with mean zero and variance given by σ_r^2.

Alternative policy rules may replace expected inflation and output with current or lagged realizations. For example, McCallum (1997) argues that a policy rule that depends on current realizations of output and inflation does not accurately depict actual information available to central banks when monetary policy decisions are made, since it takes about a full quarter to produce actual data on real GDP and price levels. He argues that the monetary policy rule should instead be expressed as a function of past data. The Taylor rule in (16) is subject to this criticism under rational expectations, but it is shown in the next section that when agents learn, expectations of future variables are completely functions of past data.

2.5 Complete Model

The complete linear New Keynesian model is represented by "IS relationship", given in equations (12) and (11); the Phillips curve in equation (14), and the Taylor rule in equation (16). These equations determine the dynamics of the output gap ($\tilde{y}_t$), the marginal utility of income gap ($\tilde{\lambda}_t$), the inflation rate (π_t), and the interest rate (r_t). The model is subject to three structural shocks: the natural rate shock, which has an autoregressive evolution given in equation (13); the cost push shock, whose evolution is given in equation (15), and the monetary policy shock.

3 Learning

The log-linearized model in the previous section can be expressed in the form,

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi v_t, \tag{17}$$

$$v_t = A v_{t-1} + \epsilon_t \tag{18}$$

where $x_t' = [\tilde{y}_t \ \tilde{\lambda}_t \ \pi_t \ \hat{r}_t]'$, $v_t' = [r_t^n \ u_t \ \epsilon_{r,t}]'$, and E_t^* denotes possibly non-rational expectations. Under rational expectations, the solution of the model has the form,

$$x_t = G x_{t-1} + H v_t, \tag{19}$$

where the elements of the matrices G and H are a function of the parameters of the model and may be determined by the method of undetermined coefficients. Under rational expectations, agents know the parameters of the model and form the expectation,

$$E_t x_{t+1} = G x_t + H E_t v_{t+1}.$$

Under learning, agents do not know the parameters of the model that make up the elements of matrices G and H. Instead, agents form expectations by estimating a linear model and using this model to make forecasts for x_{t+1}. It is popular to assume that agents know the

structure of the reduced form in equation (19), then collect data to and estimate G and H by least squares. This method is employed in this paper, but there are five important questions to consider concerning agents information sets:

1. What are the most recent observations agents have in their datasets?
2. Do agents collect data on structural shocks?
3. If so, what are the most recent observations for structural shocks?
4. Do agents estimate a constant term?
5. Do agents include as explanatory variables those associated with a column of zeros in G?

I assume agents can only collect data for variables in x_t up through the previous period. This is both a realistic and greatly simplifying assumption. While current information about interest rates are available in real life, data such as real GDP and price level released by statistical agencies such as the Bureau of Labor Statistics is typically available only months after the fact. Assuming agents have only past data greatly simplifies solving the model since x_t depends on agents' expectations. Under learning, expectations are equal to least squares forecasts, which is a non-linear function of the data agents use. Assuming x_t is not part of this data avoids the problem of solving a complex, non-linear model.

This paper explores both answers to the second question on whether data on structural shocks are available to agents. Since such data is not directly observable to an econometrician, it is quite realistic to suppose agents cannot observe this data either. When agents only have data on x_t, they form estimates for the coefficient matrix G and simply ignore the term with the structural shocks (this is appropriate since the unconditional expectation for v_t is equal to zero).

One of the goals of this paper is to identify the impact of learning on the predictions of an estimated New Keynesian model. Under rational expectations agents know current period shocks, so to isolate the effects of learning from the effects of simply assuming a more limited information set I also examine the case when agents do have data on the current period structural shocks. Since structural shocks are exogenous, there are no non-linearity issues in assuming agents have current period shocks. Moreover, equation (19) shows that

under rational expectations, assuming that agents can observe current period x_t is equivalent to assuming agents have data up to the previous period for the state vector and data up to the current period for the structural shocks. Therefore letting agents have access to data on current period shocks leads to the exact same information set under learning as rational expectations.

There is no constant term in the general form of the model, given in equation (17), or in the rational expectations solution of the model, equation (19), since all the variables in the New Keynesian model are expressed in percentage deviations from either the steady state or the flexible price outcome. However, when agents learn, they are not endowed with the values of the parameters that govern the economy, so it is unreasonable to suppose agents know the steady state of the economy. A constant term is augmented to agents regressions to capture this lack of knowledge. Agents estimate the system,

$$x_t = g + Gx_{t-1} + Hv_t,$$

or in the case when structural shocks are not observable,

$$x_t = g + Gx_{t-1}.$$

Finally, I assume that agents exclude from their datasets the variables in x_t that correspond with a column of zeros in the rational expectations solution for G. In terms of the New Keynesian model, the only variables that agents exclude is the marginal utility of income, $\tilde{\lambda}_t$. The marginal utility of income does not include any predictive power that the output gap does not, so agents exclude this from their explanatory variables in their regression. Agents do still forecast the marginal utility of income in order to make optimal consumption decisions according to the Euler equation, (11).

Let Φ_t denote the time t estimate of the all the coefficients to be estimated in the learning process. These coefficients include a vector of constants, the non-zero columns in G, and all the columns in H in the case where shocks are used as explanatory variables. Let Y_t denote the time t dependent variables used in the learning process. Since time t data is not available to agents, $Y_t = x_{t-1}$. Let X_t denote the vector of time t explanatory variables. If agents

include the stochastic shocks in their explanatory variables, $X_t' = [1\ x_{t-2}'\ v_{t-1}']$, otherwise $X_t' = [1\ x_{t-2}']$. If agents use OLS they form the estimate,

$$\Phi_t' = \left(\frac{1}{t-1}\sum_{\tau=2}^{t} X_\tau X_\tau'\right)^{-1}\left(\frac{1}{t-1}\sum_{\tau=2}^{t} X_\tau Y_\tau'\right). \tag{20}$$

The OLS estimate Φ_t can be rewritten into the convenient recursive form:

$$\Phi_t = \Phi_{t-1} + g_t(Y_t - \Phi_{t-1}X_t)X_t'R_t^{-1}, \tag{21}$$

$$R_t = R_{t-1} + g_t(X_tX_t' - R_{t-1}), \tag{22}$$

where $g_t = 1/(t-1)$ is the learning gain.[4] The recursive form demonstrates precisely how expectations are adaptive. Agents take the previous period's estimates, Φ_{t-1} and R_{t-1}, and correct them according to the residual between the previous period's forecast and the new observation. The amount of the correction depends on the learning gain. The larger is the learning gain, the more expectations respond to the latest forecast error. With OLS and infinite memory, the learning gain approaches zero as time approaches infinity, so the effect new observations have on updating the beliefs of Φ and R diminish as the number of observations already in the sample approaches infinity.

This paper instead examines the effects of constant gain learning, where the learning gain is assumed constant over time so that $g_t = g$. This type of expectations formation is appealing because unlike OLS, it allows learning to explain macroeconomic dynamics in the long run. This is a popular framework in the learning literature and is the same type of learning that Orphanides and Williams (2005b) use to explain inflation scares, Primiceri (2006) uses to explain inflation volatility in the 1970s, and Milani (2007) uses to explain macroeconomic persistence.

Constant gain learning is equivalent to estimation by weighed least squares, where the most weight is given to the most recent observation and the weights decline geometrically with age. This is a convenient framework to examine expectations when agents believe structural changes are possible. Agents do not have any information as to what types of

[4]To show this, let $R_t = \frac{1}{t-1}\sum_{\tau=2}^{t} X_\tau X_\tau'$ and $\Phi' = R_t^{-1}\left(\frac{1}{t-1}\sum_{\tau=2}^{t} X_\tau Y_\tau'\right)$

structural changes are possible, or with what probabilities. Rather, they have constant suspicion that the way the economy operates may have changed, so most recent observations are given the most weight. It has also been suggested, for example by Evans and Honkapohja (2001) and Sargent (1999), that the constant gain learning algorithm in equations (21) and (22) closely resembles expectations when agents use ordinary least squares, but with a rolling window of data where the sample size is approximately $1/g$. The constant gain learning algorithm is not identical to this scenario since it implies a weighted least squares procedure. However, the weight an additional observation under the rolling window algorithm is equal to the inverse of the sample size, which is equal to the constant learning gain.

Let $\hat{g}_{0,t}$ denote the estimated constant term in Φ_t, and let $\hat{G}_t$ and $\hat{H}_t$ denote the time t estimate of G and H, respectively, obtained from Φ_t, where $\hat{H}_t$ is simply set equal to zero in the case when structural shocks are not observable. Agents' expectation of x_{t+1} is given by,

$$E_t^* x_{t+1} = \hat{g}_{0,t} + \hat{G}_t E_t^* x_t + \hat{H} E_t v_{t+1} \tag{23}$$

Note that equation (23) assumes that expectations about future shocks, v_{t+1}, are rational. This is a common simplifying assumption made in learning models. It is possible to allow agents to also estimate the coefficients in the shock process, but the dynamics deriving from this additional complication are negligible. Since time t observations are not yet available to agents, agents must also estimate x_t by least squares. The time t estimate of x_t is given by,

$$E_t x_t^* = \hat{g}_{0,t} + \hat{G}_t x_{t-1} + \hat{H} v_t. \tag{24}$$

Plugging this into equation (23) yields,

$$E_t^* x_{t+1} = (I + \hat{G}_t)\hat{g}_{0,t} + \hat{G}_t^2 x_{t-1} + \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right) v_t. \tag{25}$$

Plugging the agents' forecast, (25), into the structural form of the model, (17), leads to the following actual law of motion for x_t,

$$x_t = \Omega_0^{-1}\Omega_2 \left(I + \hat{G}_t\right) \hat{g}_{0,t} + \Omega_0^{-1} \left(\Omega_1 + \Omega_2 \hat{G}_t^2\right) x_{t-1} + \Omega_0^{-1} \left[\Psi + \Omega_2 \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right)\right] v_t. \tag{26}$$

4 Estimation

4.1 Maximum Likelihood

The model is estimated with quarterly U.S. data from 1960:Q1 through 2007:Q1 on the output gap, as measured by the congressional budget office, the inflation rate of the consumer price index, and the federal funds rate. The model is estimated by maximum likelihood using the Kalman filter procedure described by Hamilton (1994) that maps the state equations (26) and (18) and a system of observation equations to a log-likelihood function. The observation equations are given by,

$$GAP_t = 100\tilde{y}_t,$$
$$INF_t = \pi^* + 400\pi_t,$$
$$FF_t = r^* + \pi^* + 400\hat{r}_t,$$

where GAP_t denotes data on the output gap, INF_t denotes data on the annualized quarterly inflation rate, and FF_t denotes the annualized quarterly federal funds rate. The state variables are multiplied by 100 to convert the decimals into percentages, and the inflation rate and federal funds rate are further multiplied by 4 to convert the quarterly rates to annualized rates. The New Keynesian model assumes that the steady state inflation rate is equal to zero, but since this is not likely the case in the data, the annualized steady state inflation rate, given by π^*, is estimated along with the other parameters of the model. The steady state gross real interest rate is set equal to the inverse of the discount factor; therefore $r^* = 400(1/\beta - 1)$.

The log-likelihood is maximized with respect to the learning gain, g; the New Keynesian parameters η, σ^{-1}, γ, ρ_r, ψ_y, ψ_π, ρ_n, ρ_u, σ_n, σ_u, and σ_r; and the steady state inflation rate, π^*. Instead of estimating the intertemporal elasticity of substitution, preliminary results indicated very elastic intertemporal substitution effects so it is easier to identify the inverse of this parameter. Three parameters are not estimated. The discount factor, β, is set equal to 0.9925. This corresponds to a steady state annual real interest rate of 3% which is close to the average difference between the federal funds rate and the inflation rate over the sample period. Preliminary results indicated difficulty in identifying the elasticity of labor supply, μ. The only place this parameter appears in the model is on the Phillips curve

multiplying the marginal utility of income, $\tilde{\lambda}_t$. Equation (12) shows that when carrying out this multiplication, μ and σ appear multiplicatively, causing weak identification. Therefore, the elasticity of labor supply is set equal to zero. This implies that there are no changes in labor supply decisions that effect firms' marginal costs, and therefore there are no changes in labor supply that arise from firms altering pricing decisions. Finally, the coefficient on the output gap in the Phillips curve, κ, is set equal to 0.1. Preliminary results indicated estimates of κ infinitely close to zero with a very high degree of precision, which has the unrealistic implication that prices are completely fixed for all time. Ireland (2004) reports the same difficulty and also sets $\kappa = 0.1$ prior to estimating the model.

4.2 Initial Conditions

Before estimating the model, it is necessary to specify initial conditions for the learning process given in equations (21) and (22). Unlike specifying initial conditions for the Kalman filtering procedure, the choices for initial learning matrices, Φ_0 and R_0, can have a dramatic effect on the estimation results. Despite this dependence, there is little general consensus for how initial expectations should be specified.

Williams (2003) shows that using the rational expectations solution for initial expectations produces nearly identical dynamics as assuming expectations are rational throughout the sample. Given the model is E-stable, this result is not too surprising. If the conditions for E-stability are met, under a decreasing learning gain consistent with OLS, the model will converge to the rational expectations solution when in the neighborhood of this solution. Williams shows with simulations that with a constant gain, the dynamics under learning do not significantly differ than under rational expectations.

Most initialization methods are therefore based on pre-sample evidence. Slobodyan and Wouters (2007) estimate the rational expectations version of the model on pre-sample data, and use the implied expectations as the initial condition for the sample. Milani (2007) sets initial expectations based on statistical evidence with de-meaned pre-sample data, with a few exceptions. For example he argues that agents perceived zero persistence in inflation at the beginning of the sample, when pre-sample evidence indicated it was low. Moreover, he assumes agents do observe structural shocks and so sets the initial coefficients in H equal

to zero. Primiceri (2006) calibrates initial conditions with the argument that the initial conditions are close to observed pre-sample evidence, and that the initial conditions describe well the behavior of the economy in the opening periods of the sample.

In this paper, I examine the following four specifications for how agents form expectations, and how expectations are initialized:

Case 1. Rational expectations.

Case 2. Learning with observable shocks and initial conditions set equal to rational expectations.

Case 3. Learning without observable shocks and initial conditions set equal to rational expectations.

Case 4. Learning without observable shocks and initial conditions set equal to pre-sample evidence.

Rational expectations is estimated as a baseline case for which to make comparisons. Case 2 can be viewed as the smallest step away from rational expectations. Agents have the same information set and expectations at the beginning of the sample. This implies that rational expectations is actually the special case of this learning framework where the learning gain is equal to zero. As the learning gain is estimated jointly with the other parameters of the model, the statistical significance of this parameter from zero can formally reject or fail to reject the null hypothesis that expectations are rational.

Case 3 makes another incremental step away from rational expectations. Agents again learn according to constant gain least squares, and their initial conditions for the learning matrices are equal to the rational expectations values, but agents are not able to collect data on past shocks in order to use them as explanatory variables. Due to this difference, Case 3 does not nest rational expectations.

Case 4 assumes the agents have the same information set as Case 3, but the initial conditions for the learning process matrices are different from the rational expectations solution. The initial conditions are set equal to constant gain least squares estimates from pre-sample data. Equations (21) and (22) describe the least squares learning process with any given learning gain, g_t. When the learning gain is constant, repeated substitution of

these equations can show that the learning matrices are given by,

$$R_t = \sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \tag{27}$$

$$\Phi_t = \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \right)^{-1} \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} Y'_{t-\tau} \right) \tag{28}$$

Pre-sample data on the output gap, inflation rate, and federal funds rate are collected for the period 1954:Q3 through 1959:Q4. Pre-sample data on the output gap is divided by 100 to convert it to pre-sample data for $\tilde{y}_t$. The steady state levels for the inflation rate and nominal interest rate are removed from pre-sample data on the inflation rate and federal funds rate and these are divided by 400 to be put in terms of quarterly rates in the model. The weighted least squares procedure in equations (27) and (28) is run on this pre-sample data to form matrices for Φ_0 and R_0 for the beginning sample period 1960:Q1.

5 Results

In this section I present the maximum likelihood estimation results for each of the four expectations frameworks. To determine the role learning, initial expectations, and agents information sets have on the estimation results I look at the parameter results for each model in turn. After understanding differences in parameter estimates I compare the relative fit of the models in terms of in-sample residuals and out-of-sample forecast errors. Finally I show the roles the structural shocks play on the dynamics of model by examining impulse response functions and the predicted paths of the structural shocks over the sample period.

5.1 Parameter Estimates

Case 1: Rational Expectations

Table 1 shows the parameter estimates for all four specifications. The first two columns are the results for the rational expectations model. The results show habit formation is a very strong source of output persistence, with $\eta = 0.99$. The high estimate for habit formation is similar to Smets and Wouters (2005) and (2007) estimates from a larger New Keynesian

estimated by Bayesian methods on U.S. data. This result is also consistent with Milani (2007) finding that habit formation is significant when expectations are rational. Despite the evidence for strong persistence in output, the estimated degree of price indexation is essentially equal to zero. This is in contrast with Smets and Wouters and Milani who find degrees of price indexation very close to unity. However, estimates for these degrees of persistence vary substantially across the empirical macroeconomics literature. For example, Ireland (2004) estimates a similar model by maximum likelihood and finds small degrees of persistence in both output and inflation. Nason and Smith (2005) use method of moments procedures to identify the Phillips curve and find point estimates for indexation close to 0.3. Cogley and Sbordone (2005) also use method of moments procedures and indexation is equal to 0.0.

The inverse intertemporal elasticity of substitution is $\sigma^{-1} = 0.0015$ which is very small compared to much of the macroeconomics literature. This implies that consumption decisions are very sensitive to changes in the expected real interest rate. Milani (2007) for example finds an estimate for the inverse elasticity of substitution approximately equal to 0.26 when expectations are rational. Other papers find much higher estimates. Giannoni and Woodford (2003) find the inverse elasticity approximately equal to 1.51, and Smets and Wouters (2005) find this parameter approximately equal to 1.62. Some empirical work, such as Ireland (2004), simply uses a log utility function which implicitly assumes the elasticity is equal to 1. It will be seen in the cases below that the estimate for this parameter is sensitive to the expectations framework.

Case 2: Learning with RE Initial Conditions

The next two columns of Table 1 show the parameter estimates under learning, when agents have the same information set as rational expectations and when initial expectations are set equal to the rational expectations solution. As mentioned above, rational expectations is the special case of this model where the constant learning gain is equal to zero. The estimate for the learning gain is small, $g = 0.0119$, but is statistically significantly greater than zero, which implies significant statistical evidence for learning. This learning gain corresponds to agents using approximately the last 84 observations to form their expectations,

or about 21 years of data.

Many parameters estimates are very different than under rational expectations. The degree of habit formation dropped to about $\eta = 0.65$ and the degree of inflation indexation increased dramatically from zero to $\gamma = 0.71$. This implies when agents learn, past inflation is significant in explaining forecasts for future inflation. The inverse elasticity of substitution jumped to $\sigma^{-1} = 0.42$, which is closer to other estimates found in the literature. The lower intertemporal elasticity of substitution implies that consumption decisions are less responsive to the expected real interest rate. The monetary policy parameters also indicate that dynamics in the data are less responsive to expectations under learning. The response of monetary policy to expected output and expected inflation decreased to $\psi_y = 0.08$ and $\psi_\pi = 1.74$, respectively.

Case 3: Learning with Unobservable Shocks and RE Initial Conditions

In the next learning case agents do not collect data on structural shocks. Because shocks cannot directly influence expectations, all other things remaining the same, agents' forecasts should be less volatile. The fourth and fifth columns of results in Table 1 show the parameter estimates for this framework. The learning gain is approximately, $g = 0.0202$, which is nearly twice the size as in Case 2, and given the small standard errors, the estimate is significantly higher. Since the shocks do not directly influence the volatility of expectations, the estimation results predict volatility in expectations is due to a higher learning gain. This difference in the learning gain may appear small, but when interpreting it from the viewpoint of the number of past observations agents use helps put it in perspective; in Case 3 agents use about 50 observations to form their expectations, or just over 12 years of data.

The parameter estimates for inverse intertemporal elasticity of substitution and monetary policy parameters indicate that expectations play a larger role in inflation and output determination when agents do not observe structural shocks. The inverse elasticity of substitution is $\sigma^{-1} = 0.03$ which is smaller than in Case 2, but not nearly as small as under rational expectations. Monetary policy parameters indicate stronger responses to expectations of inflation and the output gap, but still predict responses smaller than rational expectations.

Assuming learning with a limited information set therefore still leads to the conclusion

that inflation and output dynamics are less responsive under learning than under rational expectations. However, the limited information set leads to greater volatility of expectations and greater sensitivity of consumption choices and monetary policy to expectations than under learning with a full information set.

Case 4: Learning with Pre-Sample Initial Conditions

The final case uses the same limited information set as Case 3, but sets expectations for the beginning sample period equal to pre-sample weighted least squares results. The estimate for the learning gain is approximately $g = 0.0175$ which corresponds to a rolling window of over 57 observations or just over 14 years of data. Again the learning gain is statistically significantly greater than zero. This implies that expectations are adaptive over the sample period.

The estimates for the degrees of persistence are all significantly positive, but are not so close to unity. Habit formation is $\eta = 0.71$ and inflation indexation is $\gamma = 0.63$. This is in direct contrast to the Milani (2007) finding that these degrees of persistence are significant under rational expectations, but learning causes these to fall close to zero. One possible explanation for the difference in these findings is the estimation procedure. That paper uses Bayesian methods, whereas this paper uses maximum likelihood. The initial conditions for expectations are also somewhat different. Milani calibrates the initial expectations according to pre-sample estimation results from a first order vector autoregression (VAR(1)), but with some exceptions. In his paper, initial expectations for inflation persistence are set equal to zero, output gap persistence is set below pre-sample evidence, and the sensitivity of inflation to the output gap is set above pre-sample sample evidence. Moreover, the initial conditions based on pre-sample evidence for Case 4 of this paper is not set according to pre-sample VAR(1), but the pre-sample results from the weighted least squares vector autoregression given in equation (28) that is consistent with constant gain learning, for a given estimate of the learning gain.

5.2 Model Fit Comparisons

Given the different predictions of the four models, I turn to examine how well each model fits the data, and examine whether any of the learning models provides a better fit to the data

during periods of the sample that is characterized by excess volatility, as it has been proposed by some authors that learning may help explain run-ups of inflation and subsequent declines. The first three rows of Table 2 show the root mean squared residuals for each model. The results indicate a very similar performance of all four models for all three variables. The best performing model is actually the rational expectations model, but the improvement is very small.

To determine whether the learning models can explain the time varying volatility in macroeconomic activity throughout the sample, the bottom three rows of Table 2 report the autocorrelation of the square of the residuals. If a model is well specified and can explain changes in volatility in the data, then the volatility of the residuals should be constant throughout the sample and therefore the autocorrelation should equal zero. The autocorrelation of the squared residuals are small but do indicate some persistence in volatility of the residuals. The best performing model under this criteria for the output gap and federal funds rate is the learning model based on pre-sample initial conditions. The autocorrelation for the output gap in Case 4 is approximately 0.09, compared to values above 0.13 in the other frameworks. There is a small improvement in the autocorrelation for the federal funds rate, but it is still significantly above zero. The best performing model for inflation using this criteria is Case 3, learning when agents do not have data on the structural shocks. In this case the autocorrelation of the residuals is approximately 0.11 compared to values of about 0.18 and above for the other cases.

To see where the models are making their largest errors, Figure 1 shows the plots of the forecast errors over the sample period for each of models. The shaded regions indicate periods of recession as determined by the National Bureau of Economic Research. The numbers in parentheses for the learning models are the correlation with the evolution of the forecast errors predicted by the rational expectations model. Looking across the figure one can see the forecast errors for all the variables for all the learning frameworks are highly correlated with the forecast errors under rational expectations. This implies that none of the learning models can explain any better the changes in macroeconomic volatility than the rational expectations model. The largest forecast errors for the output gap are made in the recessions of the 1970s and early 1980s, the period characterized by high inflation and

relatively large macroeconomic volatility. The forecast errors for the output gap become relatively small after 1984 for every model, the period commonly referred to in the empirical monetary literature as the Great Moderation. The forecast errors for inflation similarly are largest in the middle 1970s and the early 1980s, then become relatively smaller.

The models all have similar in-sample performance, but to to determine if learning dynamics can better explain data out-of-sample, the models are re-estimated using data from 1960:Q1 through 1989:Q4 and using these sets of parameters, the models are forecast over 1990:Q1 through 2008:Q1 for long horizons. Figure 2 shows the root mean squared error of the out-of-sample forecast errors for forecast horizons of one quarter through 12 quarters. The best performing model for the output gap for forecast horizons 1 quarter through 6 quarters is the rational expectations model. However, at longer horizons, the learning model with expectations based on pre-sample data, is the best performing model. The same is not true for inflation and interest rate forecasts. For these variables, Case 4 is by far the worst performing model over the entire three year forecast horizon. For all three variables rational expectations and learning under Cases 2 and 4 have very similar out-of-sample performance over the forecast horizon.

5.3 Structural Shocks

Despite the mixed performance of the four models in in-sample and out-of-sample fit, the significance of the learning gain combined with the differences in the parameter estimates could lead to different predictions for the relative importance of the structural shocks in explaining the data. Figure 3 shows the estimated evolutions for the structural shocks, which are computed using the Kalman smoothing algorithm proposed by de Jong (1989). Again, periods of recession in the United States are shaded and the correlation of the shocks in the learning models with the rational expectations model are shown in parentheses. The correlations indicate there are some similarities in the predictions of the models, but the correlations are not quite so high as they are for the forecast errors.

The natural rate shock is highly correlated over the four models. The largest volatility of the natural rate shock is during the 1970s and early 1980s. The natural rate shock also appears responsible for the recession in 2001. The scale of the natural rate shocks show that

Case 3 predicts the largest volatility this shock. This is the case when expectations are least volatile because expectations are initialized the rational expectations solution, and structural shocks do not directly impact expectations. The cost push shock is somewhat correlated across the four models. The shock is more persistent under rational expectations, but all the models predict the largest shocks come during the periods of stagflation in the middle 1970s and early 1980s. The monetary policy shock shows that all models completely fail to deliver the change in monetary policy that occurred after Paul Volcker became chairman of the Federal Reserve. Recall, previous authors such as Primiceri (2006) and Orphanides and Williams (2005a) have suggested that the change in monetary policy was due to changes in expectations caused by learning. The results in this paper do not support such a claim. Instead the models predict very volatile monetary policy shocks start in late 1979 and become very small after 1984.

Figures 4, 5, and 6 show the predicted impulse response functions that arise from a one standard deviation shock to each of the structural shocks. Impulse responses for learning models depend on the values of the learning matrices Φ_t and R_t at the time of the impulse. The impulse response functions computed in this paper are for the state of the learning matrices at the first quarter of 2008, the last period of the sample. If the learning matrices are not equal to the rational expectations solution, then even in the absence of shocks the state variables evolve as expectations converge to the rational expectations solution. Therefore, to expose only the impact of the shocks, the plots in Figures 4, 5, and 6 show the difference between the evolution of the variables after the shock and the evolution the variables would take in the absence of any shocks.

Figure 4 shows the impulse responds functions for the natural rate shock. The results show that learning can create significant "hump-shaped" impulse responses, especially in Cases 3 and 4 when agents do not have data on structural shocks. The unit these impulse response functions are measured in is the percentage deviation of each variable from its steady state. The results indicate that Cases 3 and 4 create the largest and most prolonged effects following a shock. Because in these cases agents do not have data on structural shocks, expectations are only influenced indirectly through the effect the shocks have on the state variables. As time progresses the realizations of these state variables become data in

agents regressions, and therefore influence expectations for long periods of time. In Case 2, agents can see the structural shock and they know it is temporary, therefore a shock does not influence expectations over a long horizon.

Figure 5 shows the impulse responses due to a cost push shock. The response to inflation is very similar over all three models, with the exception of Case 4 in which inflation dies down significantly after about 6 periods, but takes a very long time to completely converge to the steady state. This finding is in contrast with Orphanides and Williams (2005b) who suggest using simulated impulse response functions from a calibrated model that an inflation shock can lead to prolonged periods of inflation. The bottom right graph in Figure 5 shows that initial expectations can lead to a very different prediction for the impact of the cost-push shock. Instead of output decreasing due to higher costs, expectations by the end of the sample are at such a point that the shock causes output to increase for a prolonged period of time, which leads to a very different path for the interest rate. This would be consistent if the increase in inflation caused by the cost push shock is instead interpreted by agents as an increase in demand.

Finally, Figure 6 shows the impulse responses from a contractionary monetary policy shock. The shape of the impulse responses in Cases 1 and 2 are very similar. However, the scale indicates that the negative impact on the output gap is much larger in Case 2 than under rational expectations. The response to the output gap in Cases 3 and 4 is much longer lived, and appears to have a oscillatory pattern after many years. Again Case 4 expectations at the end of the sample are in such a state that leads to a very different response to inflation. The contractionary monetary shock causes only a small one period decrease in the inflation rate, as the increase in interest rate causes an intertemporal substitution effect that decreases demand. The subsequent positive effect on inflation and continued negative effect on output is consistent with agents perceiving the change in interest rate as a response to a negative shock to supply.

6 Conclusion

Constant gain learning is not found to out-perform rational expectations in the context of an estimated standard Keynesian model. Previous research has suggested that constant gain learning can explain periods of prolonged inflation, run-ups of inflation and volatility and subsequent decline, and macroeconomic persistence. These claims are tested in the context of the New Keynesian model, the most popular specification for monetary models for the use of estimating and examining the impacts of monetary policy on the economy. To examine the effects of learning, initial expectations, and information sets, the rational expectations model is estimated along with three specifications for learning that differ on the assumed expectations at the beginning of the sample and whether agents can observe structural shocks.

Estimation results show that the learning gain is statistically significant in every case, indicating statistical evidence that expectations are not rational and are indeed adaptive. Moreover, the different models deliver very different parameter estimates that are responsible for the impact expectations have on consumption behavior and monetary policy. However, when comparing the models on criteria for in-sample and out-of-sample forecast errors, the rational expectations model delivered nearly as good as performance of the learning models, and compared to a learning model with initial expectations set to pre-sample evidence, the rational expectations model greatly out-perform the learning model in out-of-sample fit.

Analysis of impulse response functions showed despite the weak evidence for differences in fit, learning can have very different predictions for the effects structural shocks have on the dynamics of the model. When agents are assumed to not be able to collect data on structural shocks, the shocks produce prolonged impulse responses. Moreover, even the directions of some of the impulse response functions were shown to be quite sensitive to initial expectations.

References

Calvo, G. A. (1983): "Staggered Prices in a Utility Maximizing Framework," Journal of Monetary Economics, 12, 383–398.

Cogley, T., and A. Sbordone (2005): "A Search for a Structural Phillips Curve," Federal Reserve Bank of New Yord Staff Report no. 203.

de Jong, P. (1989): "Smoothing and Interpolation with the State-Space Model," Journal of the American Statistical Association, 84, 1085–1088.

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.

Giannoni, M. P., and M. Woodford (2003): "Optimal Inflation Targeting Rules," In: Beernanke, B.S. and M. Wordford (Eds.), Inflation Targeting. University of Chicago Press.

Hamilton, J. (1994): Time Series Analysis. Princeton University Press.

Ireland, P. N. (2004): "Technology Shocks in the New Keynesian Model," Review of Economics and Statistics, 84, 923–936.

Marcet, A., and T. J. Sargent (1989): "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," Journal of Political Economy, 6, 1306–1322.

McCallum, B. T. (1997): "Issues in the Design of Monetary Policy Rules," NBER Working Paper No. 6016.

Milani, F. (2007): "Expectations, Learning and Macroeconomic Persistence," Journal of Monetary Economics, 54, 2065–2082.

Nason, J. M., and G. W. Smith (2005): "Identifying the New Keynesian Phillips Curve," Queen's Economics Department Working Paper No. 1026.

Orphanides, A., and J. C. Williams (2005a): "Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations," Journal of Economic Dynamics and Control, 29, 1927–1950.

——— (2005b): "Inflation Scares and Forecast-Based Monetary Policy," Review of Economic Dynamics, 8, 498–527.

Preston, B. (2005): "Learning About Monetary Policy Rules when Long-Run Horizon Expectations Matter," International Journal of Central Banking, 1, 81–126.

Primiceri, G. E. (2006): "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy," Quarterly Journal of Economics, 121, 867–901.

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

Slobodyan, S., and R. Wouters (2007): "Learning in an Estimated Medium-Sized DSGE Model," Mimeo.

Smets, F., and R. Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," Journal of the European Economic Assocation, 1, 1123–1175.

——— (2005): "Comparing Shocks and Frictions in U.S. and Euro Area Business Cycles: A Bayesian DSGE Approach," Journal of Applied Econometrics, 20, 161–183.

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," American Economic Review, 97, 586–606.

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.

Table 1: Maximum Likelihood Parameter Estimates: Sample 1960:Q1 - 2008:Q1

		Case 1		Case 2		Case 3		Case 4	
Description	Parameter	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Habit Persistence	η	0.9929	0.0892	0.6515	0.0174	0.9577	0.4132	0.7065	0.2465
Inverse IES	σ^{-1}	0.0015	0.0281	0.4162	0.0536	0.0308	0.5686	0.2457	0.4541
Price Indexation	γ	0.0000	0.0377	0.7126	0.0238	0.9994	0.0754	0.6322	0.1325
MP Persistence	ρ_r	0.8857	0.0195	0.7843	0.0030	0.8558	0.0208	0.7043	0.0391
MP Output Gap	ψ_y	0.3864	0.1228	0.0758	0.0163	0.1434	0.0320	0.2265	0.0451
MP Inflation	ψ_π	3.6813	0.6479	1.7419	0.0343	2.2153	0.2974	1.5009	0.0942
Natural Rate Pers.	ρ_n	0.3636	0.0381	0.7699	0.0045	0.3060	0.0406	0.5102	0.0434
Cost Push Pers.	ρ_u	0.8568	0.0155	0.2398	0.0366	0.0000	0.0438	0.2880	0.0684
Natural Rate Std. Dev.	σ_n	0.0635	0.0128	0.0055	0.0000	0.2173	0.0584	0.0328	0.0112
Cost Push Std. Dev.	σ_u	0.0021	0.0001	0.0066	0.0003	0.0122	0.0005	0.0100	0.0008
Policy Shock Std. Dev.	σ_r	0.0032	0.0001	0.0031	0.0001	0.0030	0.0000	0.0031	0.0001
Steady State Inlfation	π^*	3.5304	0.2017	3.2369	0.3026	4.5971	0.4958	4.0443	0.3728
Learning Gain	g	–	–	0.0119	0.0015	0.0202	0.0020	0.0175	0.0027

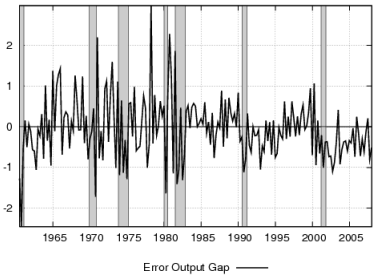
Table 2: Model Fit Comparisons

	Root Mean Squared Error			
	Case 1	Case 2	Case 3	Case 4
Output Gap	0.7554	0.7885	0.7793	0.7903
Inflation	2.3967	2.5560	2.4324	2.4373
Federal Funds Rate	1.2497	1.2557	1.2294	1.2744
	Autocorrelation Squared Error			
	Case 1	Case 2	Case 3	Case 4
Output Gap	0.1379	0.1771	0.1345	0.0867
Inflation	0.2098	0.1792	0.1145	0.2454
Federal Funds Rate	0.3225	0.4033	0.3386	0.2602

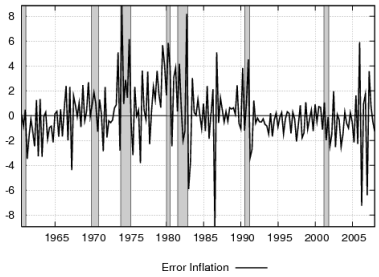
Figure 1: Forecast Errors

Case 1: Rational Expectations

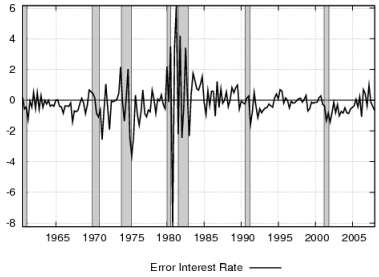
Output Gap



Inflation

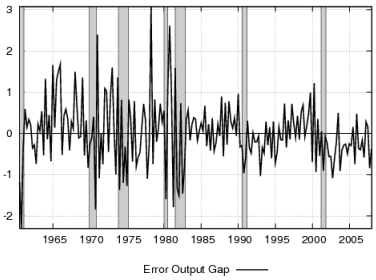


Fed Funds

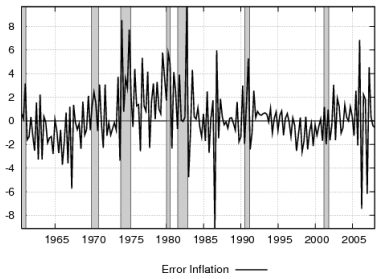


Case 2: Learing with RE Initial Conditions

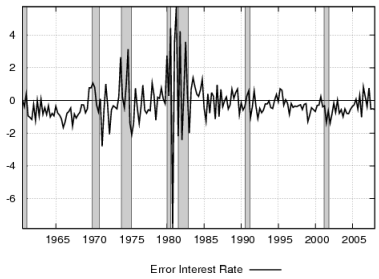
Output Gap (0.9777)



Inflation (0.9271)

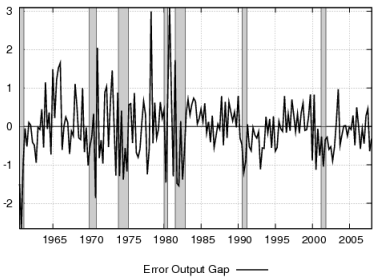


Fed Funds (0.9486)

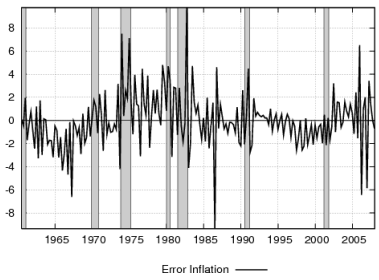


Case 3: Learing with RE Initial Conditions, Shocks Unobservable

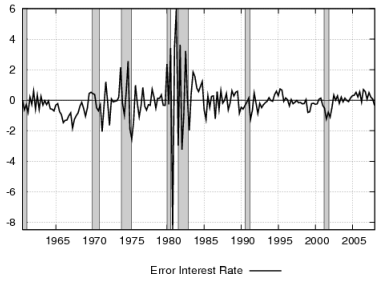
Output Gap (0.9687)



Inflation (0.9027)

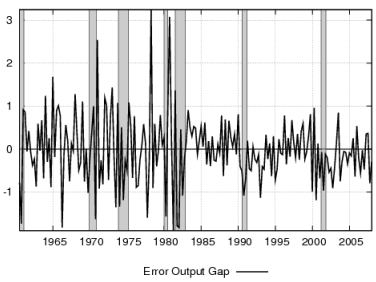


Fed Funds (0.9521)

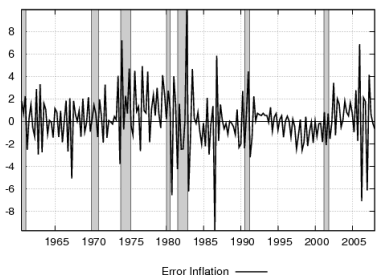


Case 4: Learning with Unobservable Shocks and Pre-Sample Initial Conditions

Output Gap (0.9237)



Inflation (0.8876)



Fed Funds (0.9084)

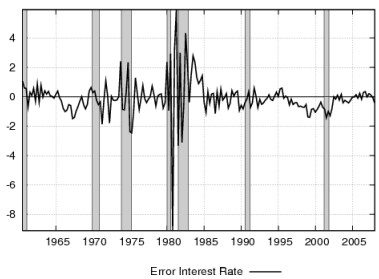
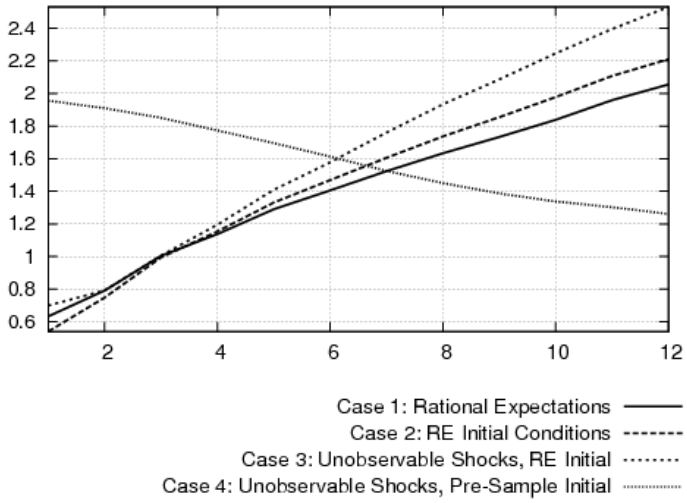
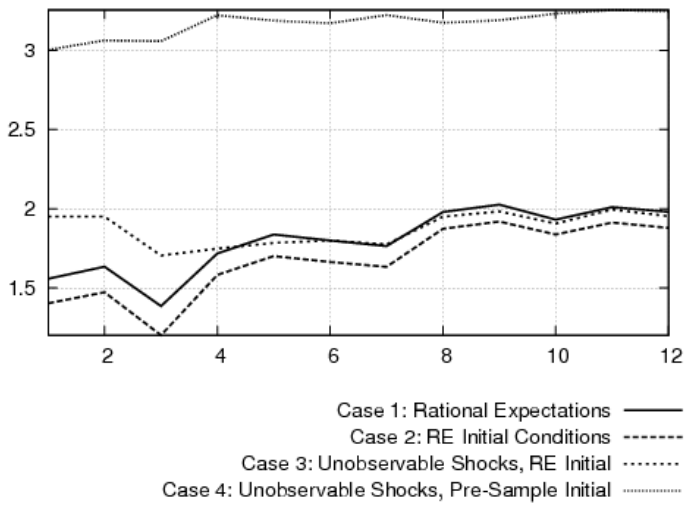


Figure 2: Out of Sample Multiperiod Forecast Errors

Output Gap



Inflation



Federal Funds Rate

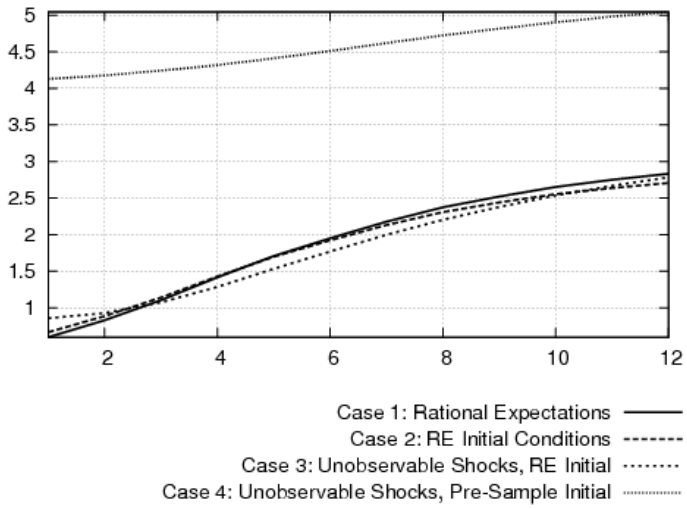
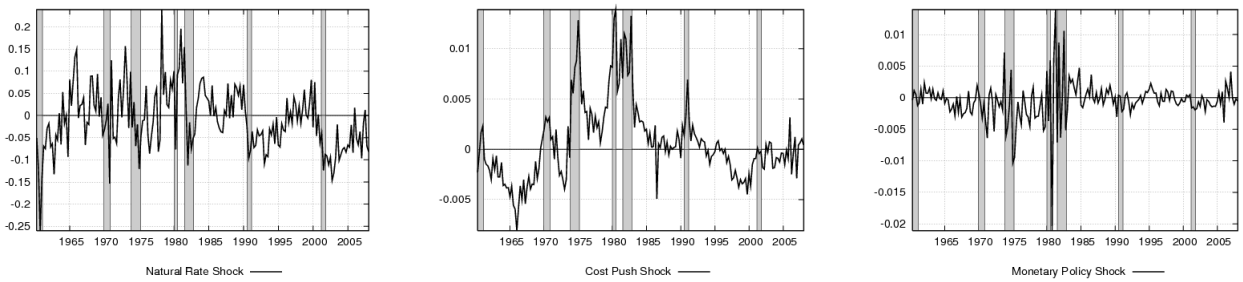


Figure 3: Smoothed Estimates of Structural Shocks

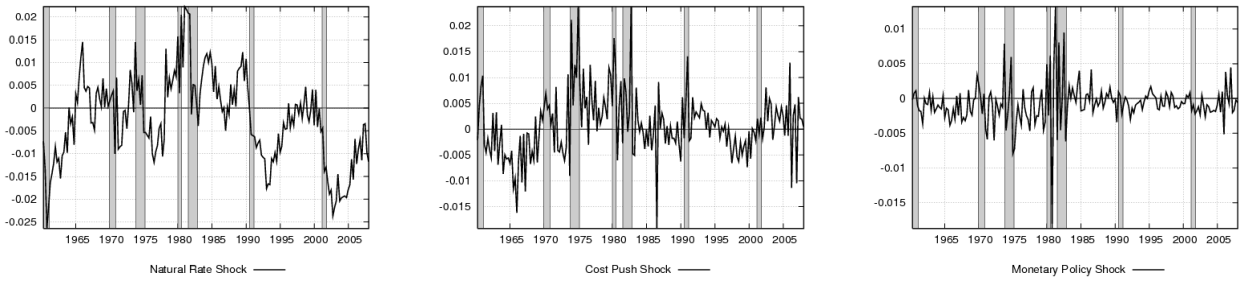
Case 1: Rational Expectations

Natural Rate | Cost-Push | Monetary Policy



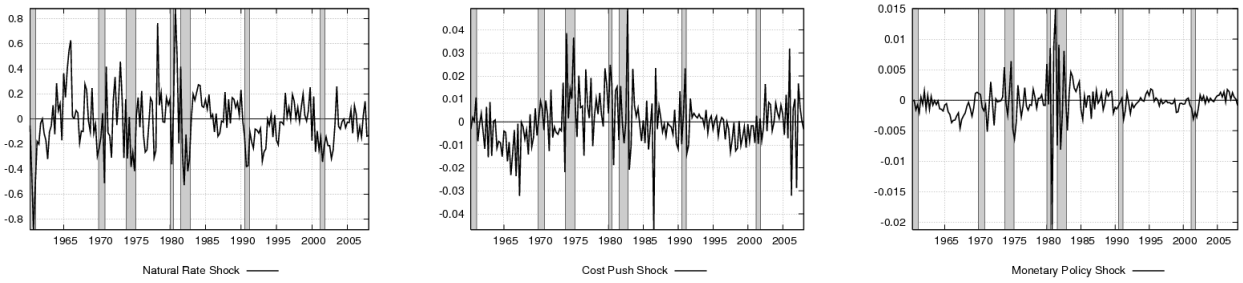
Case 2: Learning with RE Initial Conditions

Natural Rate (0.8334) | Cost-Push (0.7733) | Monetary Policy (0.9590)



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Natural Rate (0.8754) | Cost-Push (0.6215) | Monetary Policy (0.8371)



Case 4: Learning with Unobservable Shocks and Pre-Sample Initial Conditions

Natural Rate (0.8444) | Cost-Push (0.4461) | Monetary Policy (0.8271)

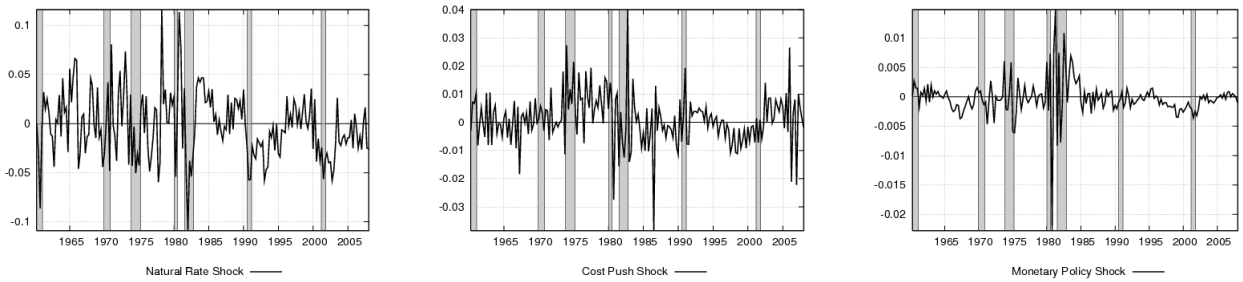
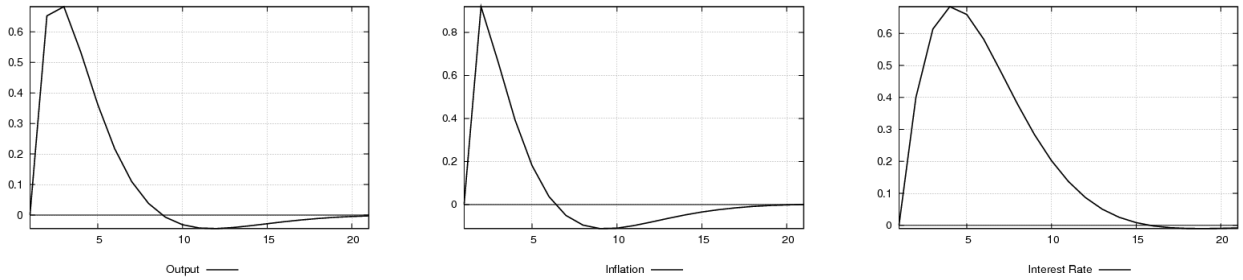


Figure 4: Natural Rate Shock Impulse Responses

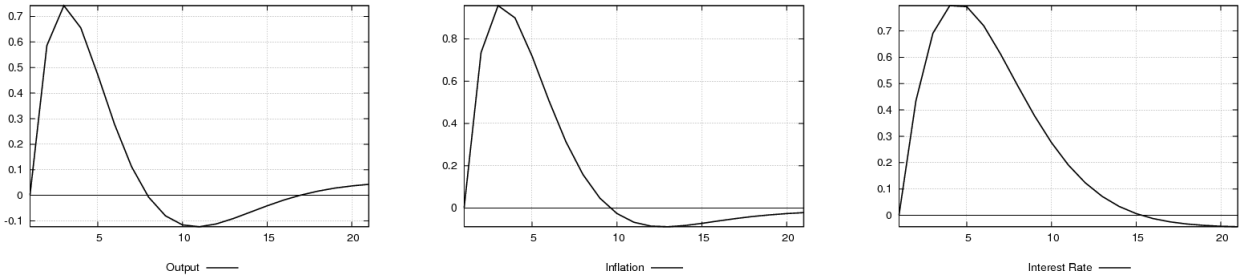
Case 1: Rational Expectations

Output Gap Inflation Interest Rate



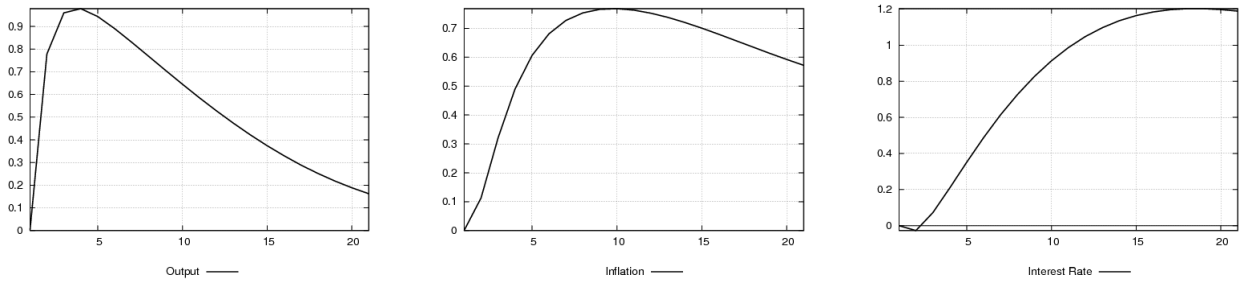
Case 2: Learning with RE Initial Conditions

Output Gap Inflation Interest Rate



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Output Gap Inflation Interest Rate



Case 4: Learning with Unobservable Shocks and Pre-Sample Initial Conditions

Output Gap Inflation Interest Rate

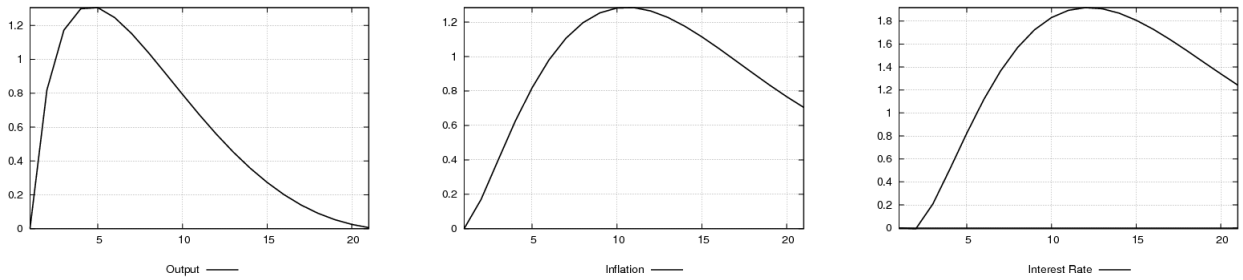
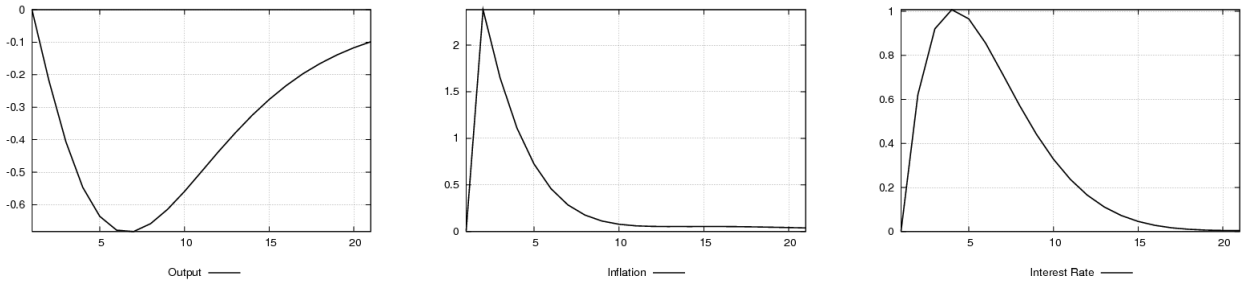


Figure 5: Cost-Push Shock Impulse Responses

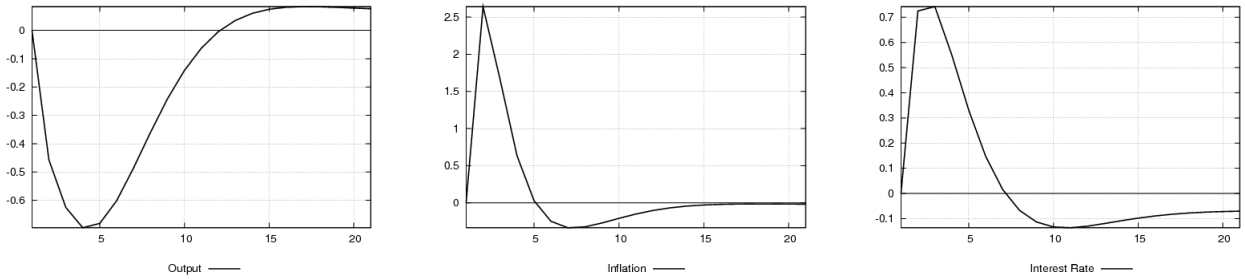
Case 1: Rational Expectations

Output Gap | Inflation | Interest Rate



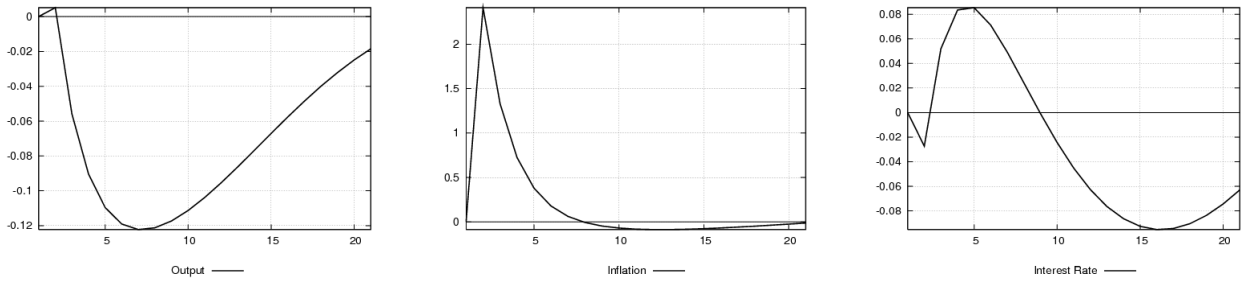
Case 2: Learning with RE Initial Conditions

Output Gap | Inflation | Interest Rate



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Output Gap | Inflation | Interest Rate



Case 4: Learning with Unobservable Shocks and Pre-Sample Initial Conditions

Output Gap | Inflation | Interest Rate

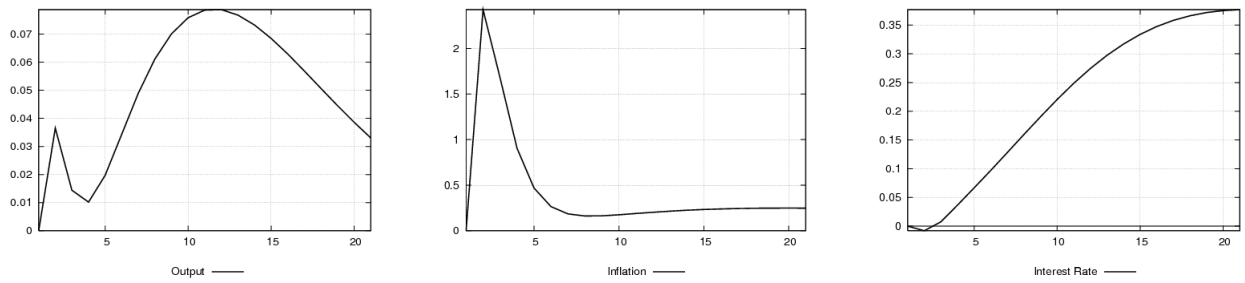
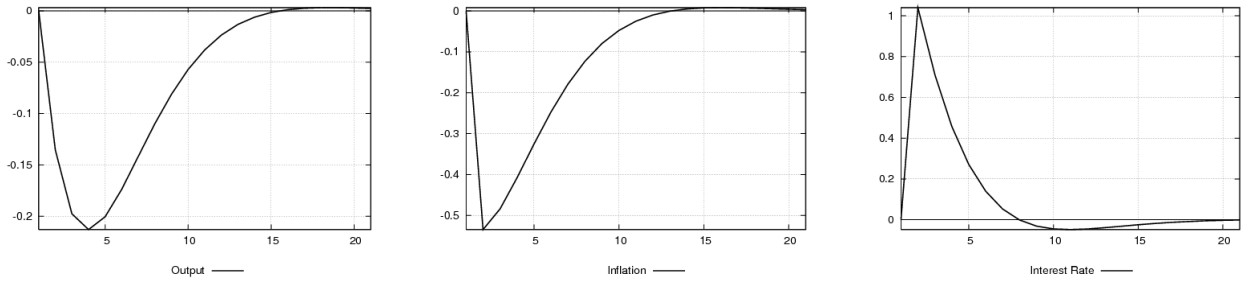


Figure 6: Monetary Policy Shock Impulse Responses

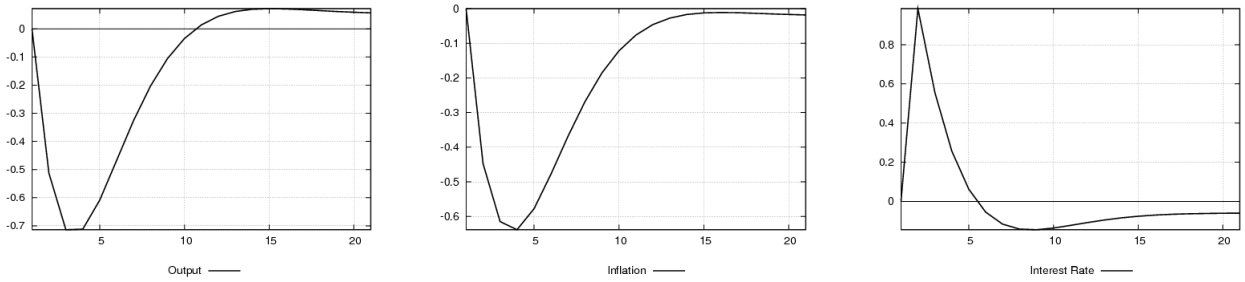
Case 1: Rational Expectations

Output Gap | Inflation | Interest Rate



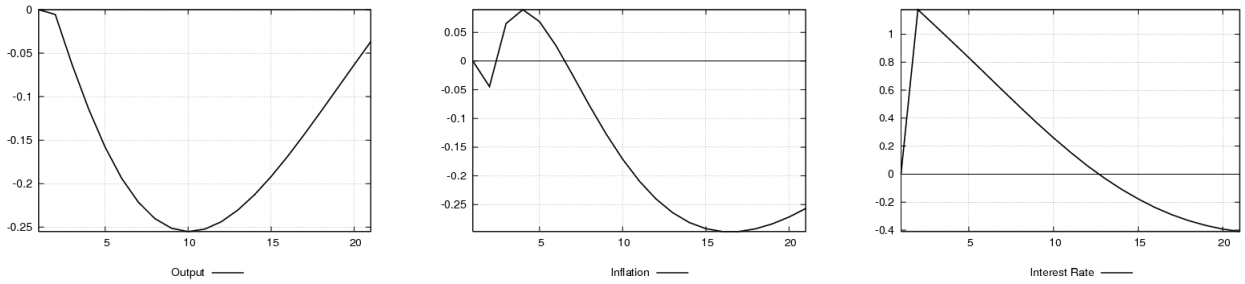
Case 2: Learning with RE Initial Conditions

Output Gap | Inflation | Interest Rate



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Output Gap | Inflation | Interest Rate



Case 4: Learning with Unobservable Shocks and Pre-Sample Initial Conditions

Output Gap | Inflation | Interest Rate