THREE ESSAYS IN ADAPTIVE EXPECTATIONS IN NEW KEYNESIAN MONETARY ECONOMIES

James M. Murray

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Economics,
Indiana University
September 2008

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee:

Eric M. Leeper, Chair

Kim P. Huynh

Brian M. Peterson

Todd B. Walker

Date of Oral Examination: July 8, 2008

Acknowledgements: I am grateful for the advice and guidance of my dissertation committee: Eric Leeper, Kim Huynh, Brian Peterson, and Todd Walker; for useful conversations with Pedro Falcão de Araujo, James Bullard, Troy Davig, Kenneth Kasa, Fabio Milani, Michael Plante, and Bruce Preston; and for comments by the participants of the 2007 Federal Reserve Bank of St. Louis Learning Week Conference, the 2007 Missouri Economics Conference, and Indiana University economics department seminars. All errors are my own.

James M. Murray

THREE ESSAYS IN ADAPTIVE EXPECTATIONS IN NEW KEYNESIAN MONETARY ECONOMIES

Abstract: This dissertation explores the empirical significance of least squares learning in estimated New Keynesian monetary models with U.S. data. Specifically, the papers set out to determine what impact learning has on dynamics of output, consumption, investment, inflation and monetary policy, and whether learning can explain empirical puzzles in the monetary literature such as the Great Moderation. In the first dissertation paper, a standard New Keynesian model is estimated with constant gain learning with three specifications for how agents' expectations are initialized. The results indicate differences in the model's prediction depending on the type of initial expectations, but the learning models do not significantly explain the data better than rational expectations. The second paper examines an extension of the model that allows for firm-specific capital accumulation. The results show that learning can lead to very different predictions for the impacts of structural shocks, depending on the choice for agents' initial beliefs. Again, constant gain learning is shown to not explain the Great Moderation any better than rational expectations. The final paper examines an extension to the learning process, where the learning gain changes endogenously with agents forecast errors. This learning framework is estimated jointly with a regime-switching volatility mechanism to determine if dynamic gain learning can lead to lower estimates for exogenously changing volatility. The results show, rather, that learning gain dynamics are quite small and are again not capable of explaining time-varying macroeconomic volatility.

Contents

List of Tables

List of Figures

Chapter 1

Initial Expectations in New Keynesian Models with Learning

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.

1.1 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 errors, 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.

1.2 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. In the next section the rational expectations are replaced by expectations under learning.[1]

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.

1.2.1 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

[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 the model.

purchases of real government bonds, $b_t(i)$, to maximize lifetime utility,

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-\frac{1}{\sigma}} \xi_t \left(c_t - \eta c_{t-1}\right)^{1-\frac{1}{\sigma}} - \frac{1}{1+\frac{1}{\mu}} n_t(i)^{1+\frac{1}{\mu}} \right], \tag{1.1}$$

subject to the budget constraint,

$$c_t + b_t(i) = \frac{1+r_{t-1}}{1+\pi_t} b_{t-1}(i) + \frac{w_t(i)}{p_t} n_t(i) + \Pi_t - \tau_t. \tag{1.2}$$

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,

$$\hat{\lambda}_t = E_t \hat{\lambda}_{t+1} + \hat{r}_t - E_t \pi_{t+1}, \tag{1.3}$$

where $\hat{\lambda}_t$ is the percentage deviation from the steady state of the Lagrange multiplier on the budget constraint, (1.2), 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,

$$\hat{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\hat{c}_{t+1} - \sigma(1+\beta\eta^2)\hat{c}_t + \sigma\eta\hat{c}_{t-1}\right] + \left(\hat{\xi}_t - \beta\eta E_t\hat{\xi}_{t+1}\right). \quad (1.4)$$

The marginal utility of income, (1.4), and the Euler equation, (1.3), make up the IS sector of the model.

1.2.2 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,

$$y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di\right]^{\frac{\theta}{\theta-1}}, \quad (1.5)$$

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,

$$y_t(i) = \left[\frac{p_t(i)}{p_t}\right]^{-\theta} y_t, \quad (1.6)$$

where $p_t(i)$ is the price of intermediate good i and p_t is the price of the final good. Substituting equation (1.6) into equation (1.5) leads to the following expression for the price of the final good in terms of the prices of intermediate goods,

$$p_t = \left[\int_0^1 p_t(i)^{1-\theta} di\right]^{\frac{1}{1-\theta}}. \quad (1.7)$$

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

[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.

choices for labor demand and labor market clearing leads to the following aggregate log-linear marginal cost,

$$\hat{\psi}_t = \frac{1}{\mu}\hat{y}_t - \hat{\lambda}_t - \left(\frac{1}{\mu} + 1\right)\hat{z}_t. \tag{1.8}$$

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:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_t(i)\pi^*_{t+T}}{p_{t+T}} \right) y_{t+T}(i) - \Psi\left[y_{t+T}(i)\right] \right\}, \tag{1.9}$$

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

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{\mu(1-\omega)(1-\omega\beta)}{\omega(\mu+\theta)}\hat{\psi}_t\right]. \tag{1.10}$$

1.2.3 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 $\tilde{y}_t = \hat{y}_t - \hat{y}_t^f$ and $\tilde{\lambda}_t = \hat{\lambda}_t - \hat{\lambda}_t^f$ 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, (1.3), and marginal utility of income, (1.4), still hold. Using these conditions and imposing goods market clearing that consumption is equal to output implies,

$$\tilde{\lambda}_t = E_t\tilde{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1} - r_t^n, \tag{1.11}$$

$$\tilde{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\tilde{y}_{t+1} - \sigma(1+\beta\eta^2)\tilde{y}_t + \sigma\eta\tilde{y}_{t-1}\right], \tag{1.12}$$

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,

$$r_t^n = \rho_n r_{t-1}^n + \epsilon_{n,t}, \tag{1.13}$$

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

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

flexible prices, equation (1.8) 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, (1.8). Plugging this expression for marginal cost into equation (1.10) 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,

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \kappa(\tilde{y}_t - \mu\tilde{\lambda}_t) + u_t\right], \tag{1.14}$$

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

$$u_t = \rho_u u_{t-1} + \epsilon_{u,t}, \tag{1.15}$$

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

1.2.4 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,

$$\hat{r}_t = \rho_r\hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi E_t\pi_{t+1} + \psi_y E_t\tilde{y}_{t+1}\right) + \epsilon_{r,t} \tag{1.16}$$

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 (1.16) 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.

1.2.5 Complete Model

The complete linear New Keynesian model is represented by "IS relationship", given in equations (1.12) and (1.11); the Phillips curve in equation (1.14), and the Taylor rule in equation (1.16). 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 ($\hat{r}_t$). The model is subject to three structural shocks: the natural rate shock, which has an autoregressive evolution given in equation (1.13); the cost push shock, whose evolution is given in equation (1.15), and the monetary policy shock.

1.3 Learning

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

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi v_t \tag{1.17}$$

$$v_t = Av_{t-1} + \epsilon_t, \tag{1.18}$$

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,

$$x_t = Gx_{t-1} + Hv_t, \tag{1.19}$$

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,

$$E_t x_{t+1} = Gx_t + HE_t v_{t+1}.$$

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 (1.19), 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 (1.19) 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 (1.17), or in the rational expectations solution of the model, equation (1.19), 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 variable 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, (1.11).

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,

$$\Phi_t' = \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau X_\tau' \right)^{-1} \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau Y_\tau' \right). \tag{1.20}$$

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

$$\Phi_t = \Phi_{t-1} + g_t (Y_t - \Phi_{t-1} X_t) X_t' R_t^{-1}, \tag{1.21}$$

$$R_t = R_{t-1} + g_t (X_t X_t' - R_{t-1}), \tag{1.22}$$

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 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 (1.21) and (1.22) 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

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

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,

$$E_t^* x_{t+1} = \hat{g}_{0,t} + \hat{G}_t E_t^* x_t + \hat{H} E_t v_{t+1} \tag{1.23}$$

Note that equation (1.23) 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,

$$E_t x_t^* = \hat{g}_{0,t} + \hat{G}_t x_{t-1} + \hat{H} v_t. \tag{1.24}$$

Plugging this into equation (1.23) yields,

$$E_t^* x_{t+1} = (I + \hat{G}_t)\hat{g}_{0,t} + \hat{G}_t^2 x_{t-1} + \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right) v_t. \tag{1.25}$$

Plugging the agents' forecast, (1.25), into the structural form of the model, (1.17), leads to the following actual law of motion for x_t,

$$x_t = \Omega_0^{-1}\Omega_2 \left(I + \hat{G}_t\right) \hat{g}_{0,t} + \Omega_0^{-1} \left(\Omega_1 + \Omega_2 \hat{G}_t^2\right) x_{t-1} + \Omega_0^{-1} \left[\Psi + \Omega_2 \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right)\right] v_t. \tag{1.26}$$

1.4 Estimation

1.4.1 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 (1.26) and (1.18) and a system of observation equations to a log-likelihood function. The observation equations are given by,

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

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 (1.12) 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 (2004b) reports the same difficulty and also sets $\kappa = 0.1$ prior to estimating the model.

1.4.2 Initial Conditions

Before estimating the model, it is necessary to specify initial conditions for the learning process given in equations (1.21) and (1.22). 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 (2005) 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 (1.21) and (1.22) 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,

$$R_t = \sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \tag{1.27}$$

$$\Phi_t = \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \right)^{-1} \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} Y'_{t-\tau} \right) \tag{1.28}$$

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 (1.27) and (1.28) is run on this pre-sample data to form matrices for Φ_0 and R_0 for the beginning sample period 1960:Q1.

1.5 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.

1.5.1 Parameter Estimates

Case 1: Rational Expectations

Table 1.1 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 (2004b) 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 (2004b), 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 1.1 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 1.1 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 (1.28) that is consistent with constant gain learning, for a given estimate of the learning gain.

1.5.2 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 1.2 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 1.2 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 1.1 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 1.2 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 3 have very similar out-of-sample performance over the forecast horizon.

1.5.3 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 1.3 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 to 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 starting in late 1979 and become very small after 1984.

Figures 1.4, 1.5, and 1.6 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 1.4, 1.5, and 1.6 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 1.4 shows the impulse responds functions for the natural rate shock. The unit these are measured in is the percentage deviation of each variable from its steady state. 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 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 shocks can 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 1.5 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 row of graphs in Figure 1.5 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 1.6 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.

1.6 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

Bullard, J., and J. Duffy (2004): "Learning and Structural Change in Macroeconomic Data," Federal Reserve Bank of St. Louis Working Paper 2004-016A.

Bullard, J., and S. Eusepi (2005): "Did the Great Inflation Occur Despite Poicymaker Commitment to a Taylor Rule?," Review of Economic Dynamics, 8, 324–359.

Bullard, J., G. W. Evans, and S. Honkapohja (2005): "Near Rational Exuberance," Federal Reserve Bank of St. Louis Working Paper 2004-025B.

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

Carceles-Poveda, A., and C. Giannitsarou (2005): "Adaptive Learning in Practice," Working Paper.

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.

——— (2008): "Expectations, Learning and Monetary Policy: An Overview of Recent Research," Centre for Dynamic Macroeconomic Analysis Working Paper CDMA08/02.

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. (2004a): "A method for taking models to the data," Journal of Economic Dynamics and Control, 28, 1205–1226.

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

——— (2005): "Irrational Expectations and Econometric Practice," Federal Reserve Bank of Atlanta Working Paper 2003-22.

——— (2006): "Changes in the Federal Reserve's Inflation Target: Causes and Consequences," Working Paper.

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 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.

Mark, N. C. (2005): "Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dyanamics," NBER Working Paper 11061.

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

Milani, F. (2005): "Learning, Monetary Policy Rules, and Macroeconomic Stability," Working Paper.

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

Murray, J. M. (2008): "Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital," Indiana University mimeo.

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.

Roberts, J. M. (1995): "New Keynesian Economics and the Phillips Curve," Journal of Money, Credit and Banking, 27, 975–984.

Rotemberg, J. (1982): "Sticky Prices in the United States," Journal of Political Economy, 90, 1187–1211.

Rotemberg, J., and M. Woodford (1997): "An Optimization Based Econometric Framework for the Evaluation of Monetary Policy," In: Bernanke, B.S. and J. Rotemberg (Eds.), NBER Macroeconomics Annual. MIT Press.

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

Sims, C. (2000): "Solving Linear Rational Expectations Models," Unpublished manuscript.

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. (2005): "Adaptive Learning and Business Cycles," Working paper.

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

Table 1.1: 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

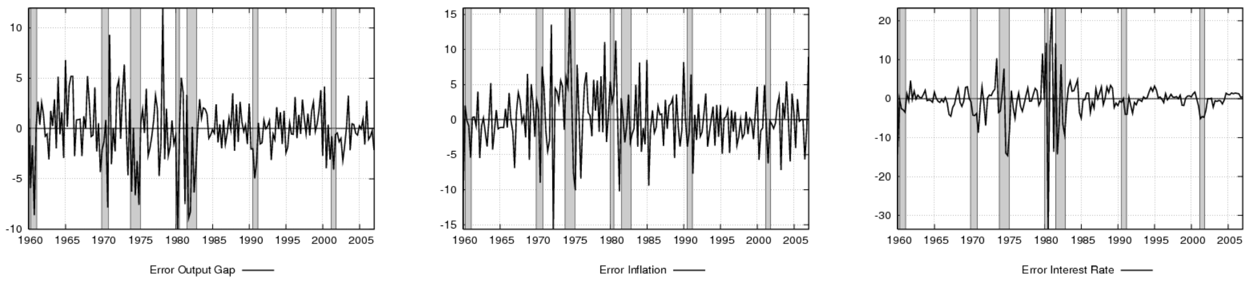
Table 1.2: Model Fit Comparisons

	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 1.1: Forecast Errors

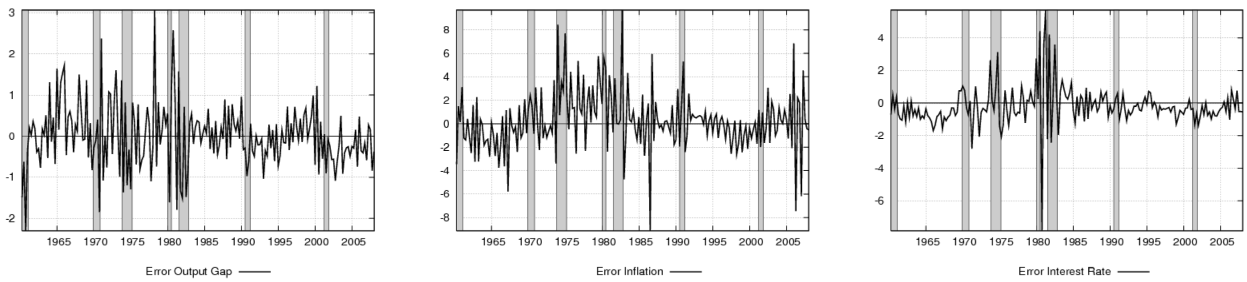
Case 1: Rational Expectations

Output Gap | Inflation | Fed Funds



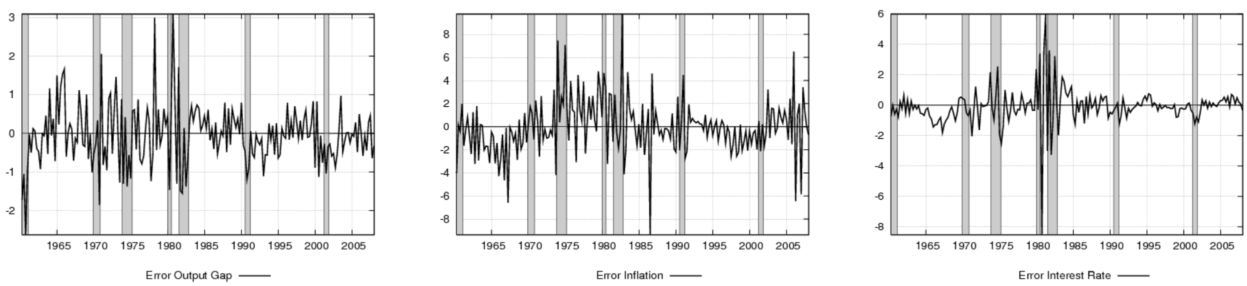
Case 2: Learing with RE Initial Conditions

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)

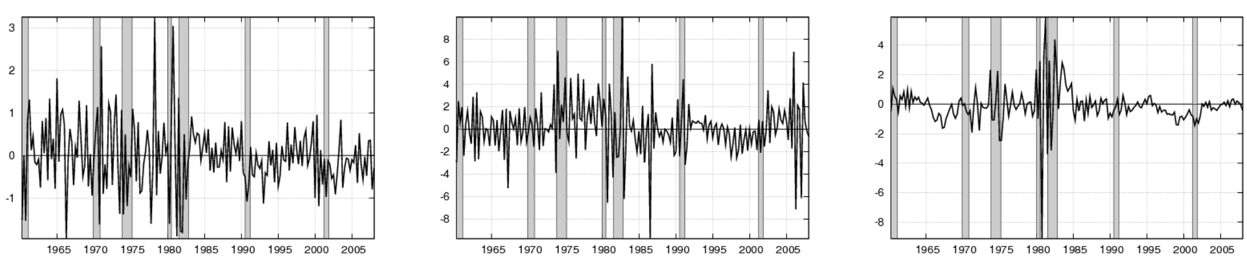
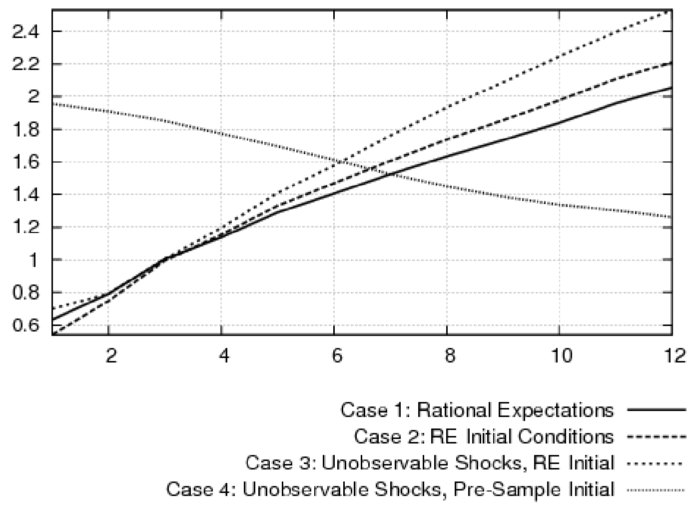
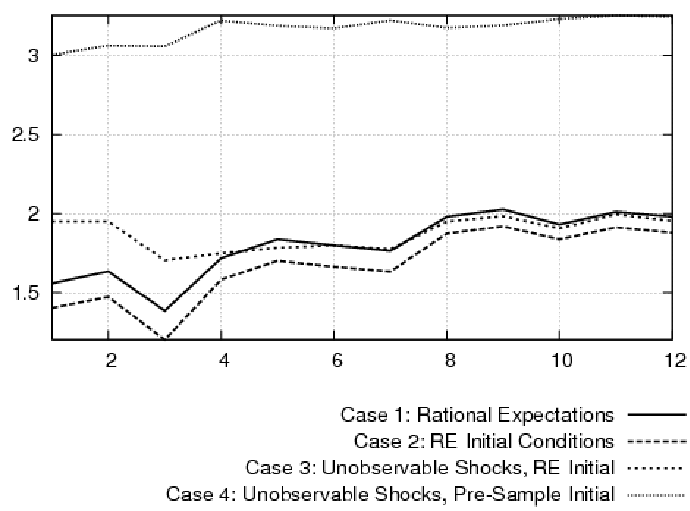


Figure 1.2: Out of Sample Multiperiod Forecast Errors

Output Gap



Inflation



Federal Funds Rate

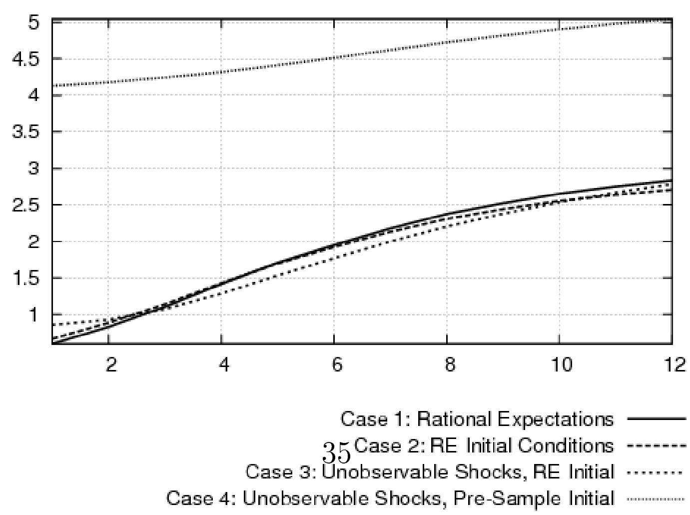
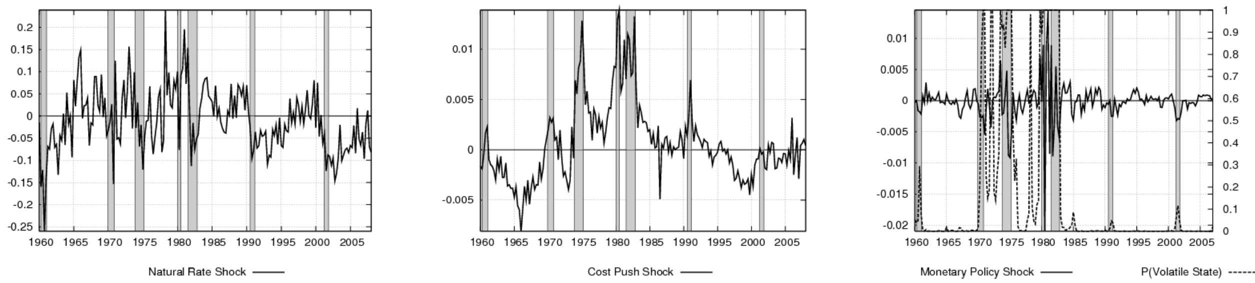


Figure 1.3: Smoothed Estimates of Structural Shocks

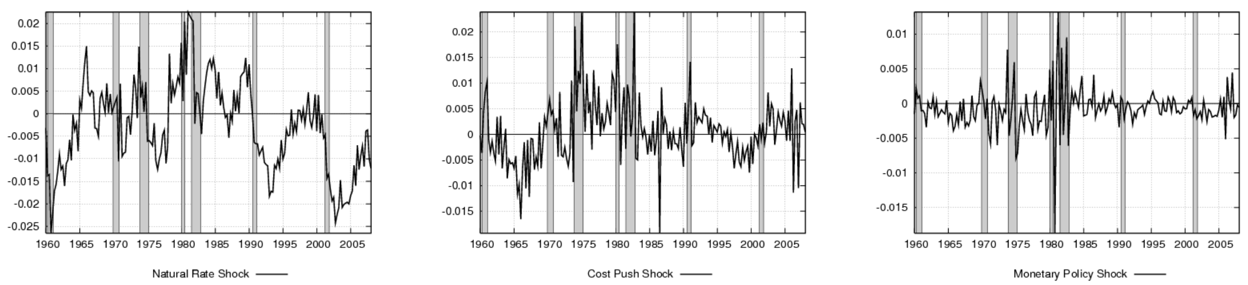
Case 1: Rational Expectations

Natural Rate | Cost-Push | Monetary Policy



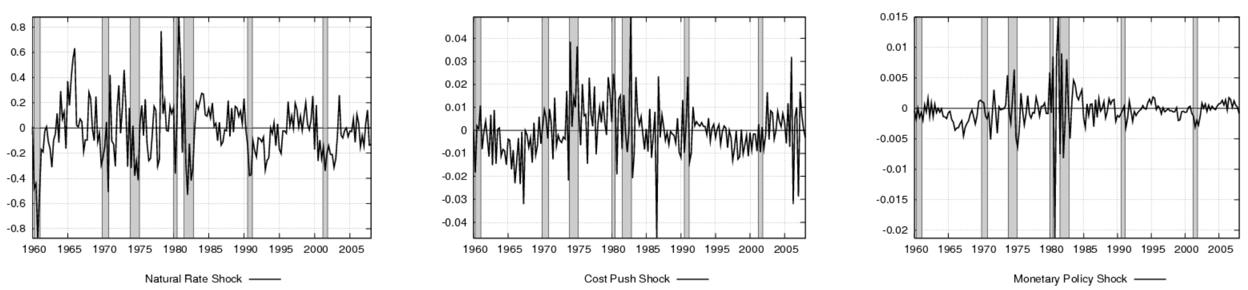
Case 2: Learning with RE Initial Conditions

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)

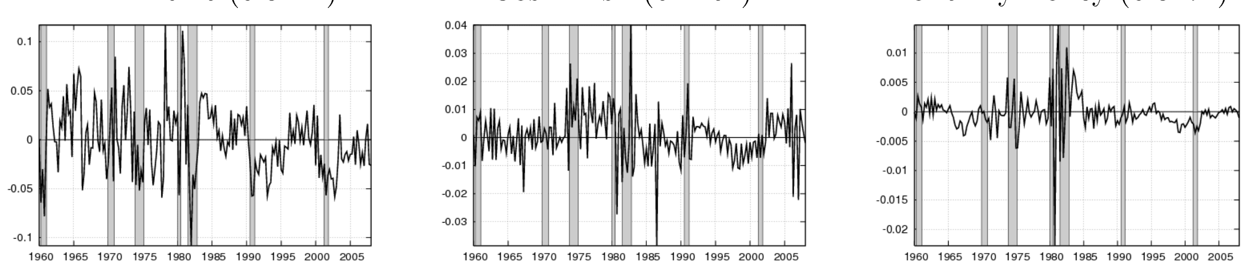
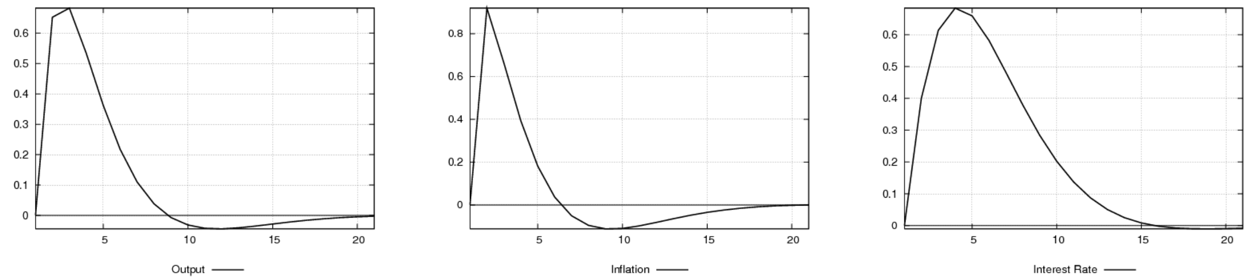


Figure 1.4: Natural Rate Shock Impulse Responses

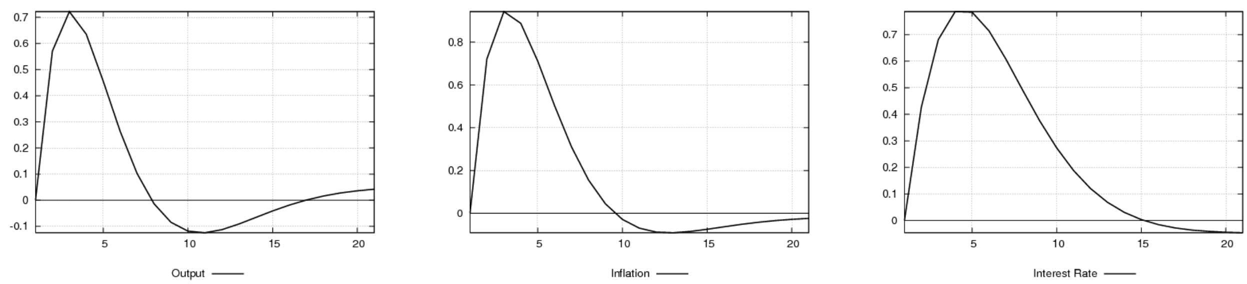
Case 1: Rational Expectations

Output Gap | Inflation | Interest Rate



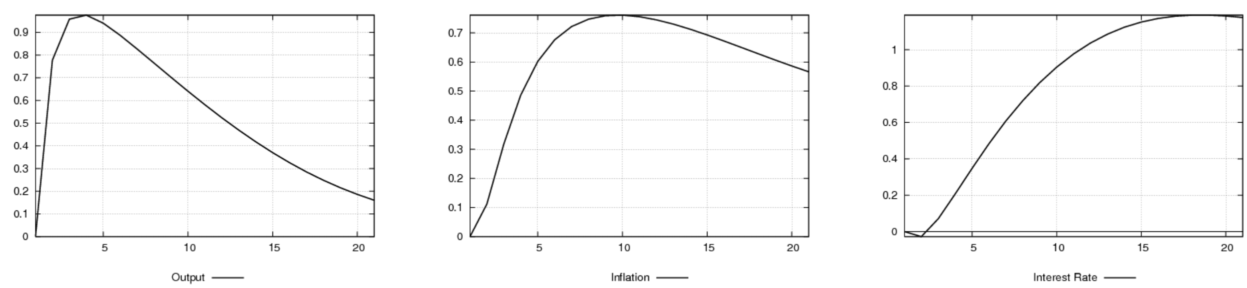
Case 2: Learning with RE Initial Conditions

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

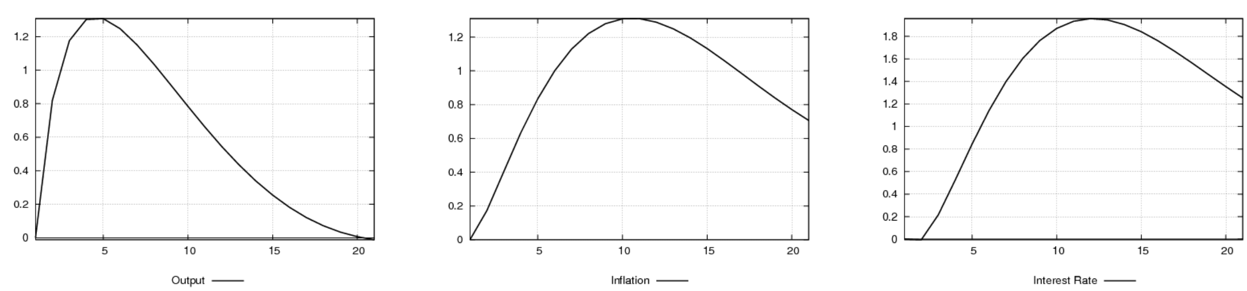
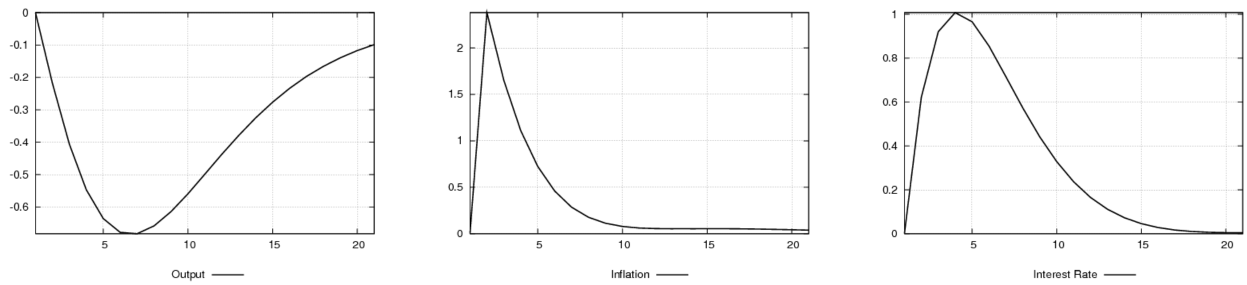


Figure 1.5: Cost-Push Shock Impulse Responses

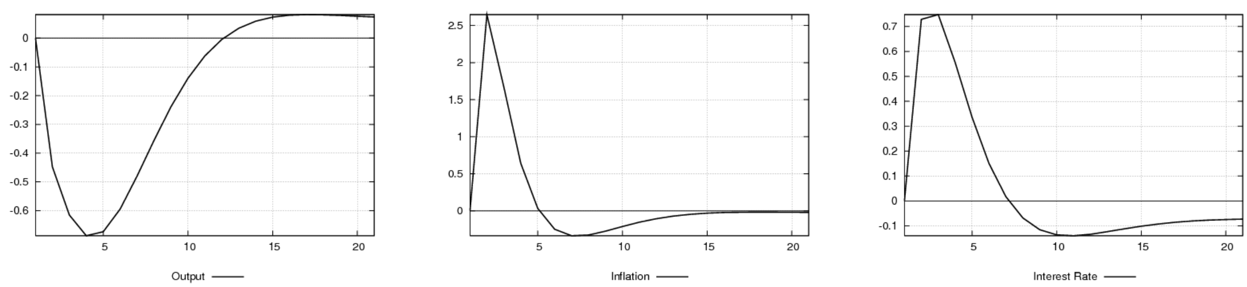
Case 1: Rational Expectations

Output Gap | Inflation | Interest Rate



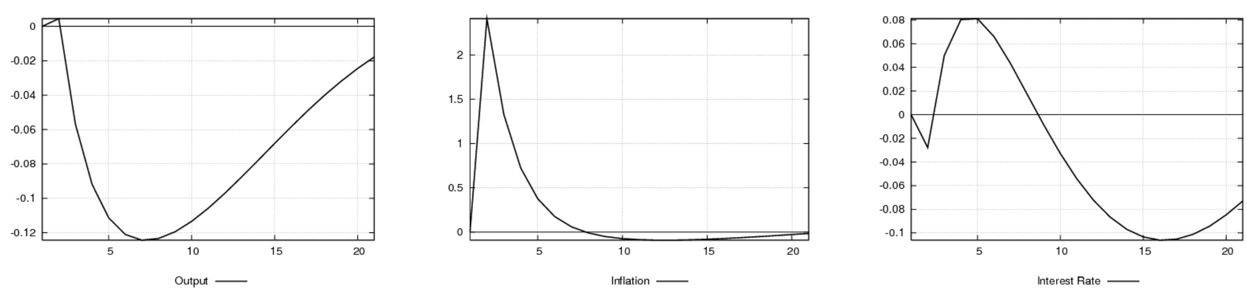
Case 2: Learning with RE Initial Conditions

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

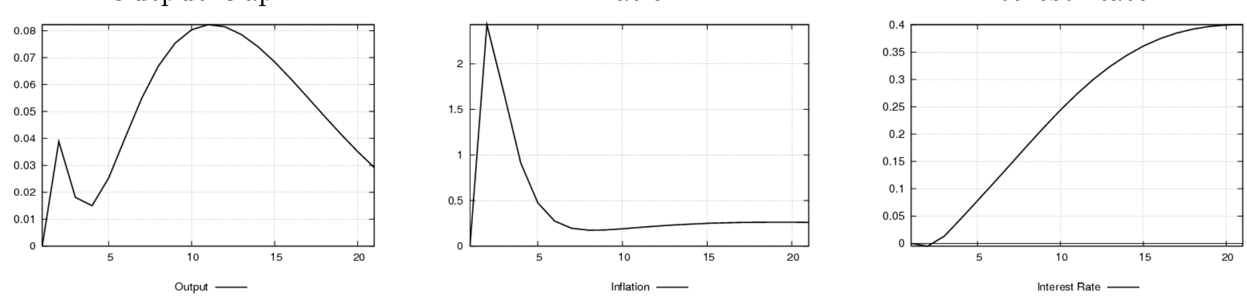
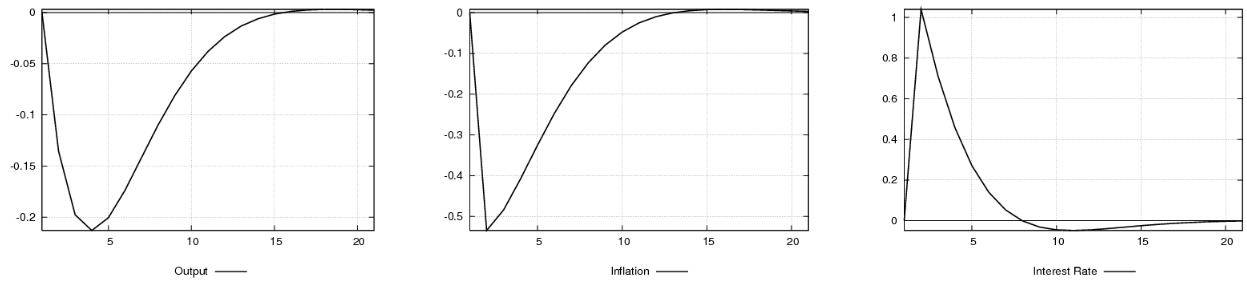


Figure 1.6: Monetary Policy Shock Impulse Responses

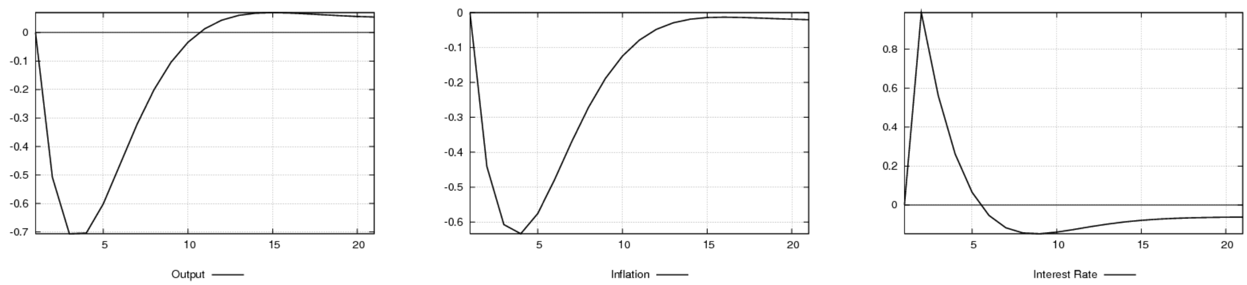
Case 1: Rational Expectations

Output Gap | Inflation | Interest Rate



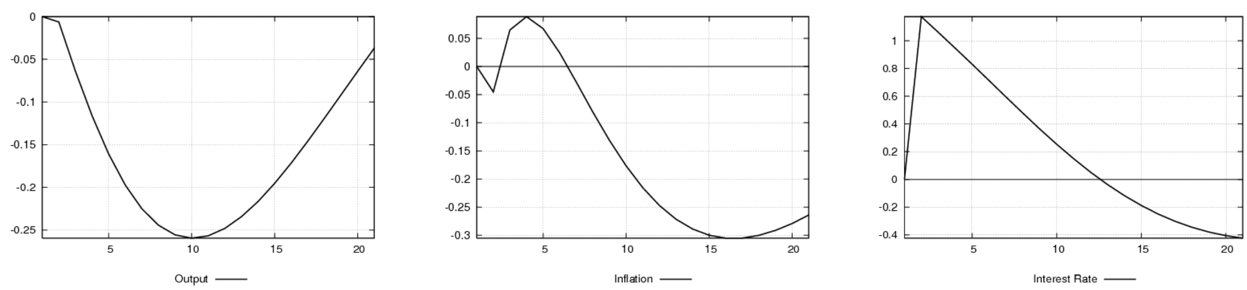
Case 2: Learning with RE Initial Conditions

Output Gap | Inflation | Interest Rate



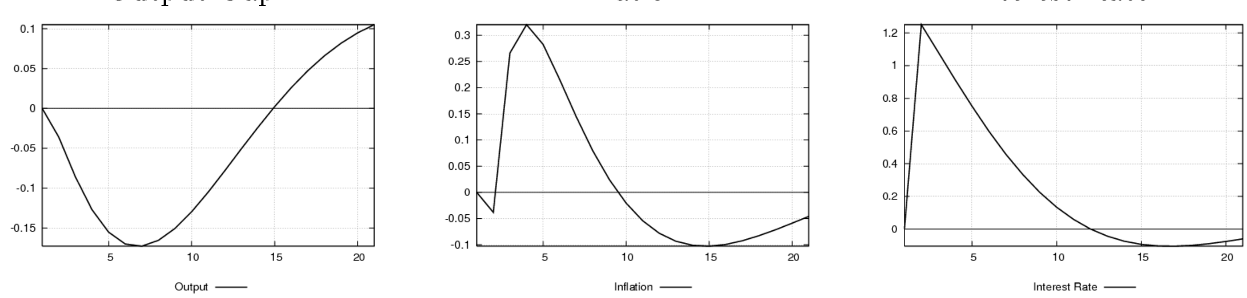
Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Output Gap | Inflation | Interest Rate



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

Output Gap | Inflation | Interest Rate



Chapter 2

Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital

Abstract: This paper examines the empirical significance of learning, a type of adaptive, boundedly rational expectations, in the U.S. economy within the framework of the New Keynesian model with endogenous capital accumulation. Estimation results for learning models can be sensitive to the choice for agents' initial expectations, so three methods for choosing initial expectations are examined. Maximum likelihood results show that learning under all methods do not significantly improve the fit the model. The evolution of forecast errors show that the learning models do not out perform the rational expectations model during the run-up of inflation in the 1970s and the subsequent decline in the 1980s, a period of U.S. history which others have suggested learning may play a role. Despite the failure of learning models to better explain the data, analysis of the impulse response functions and paths of structural shocks during the sample show that learning can lead to different explanations for the data.

2.1 Introduction

Rational expectations is one of the most common assumptions in dynamic macroeconomic models. While it is usually made for mathematical convenience, the assumption regarding expectations formation can have non-trivial effects on a model's dynamics. In particular, a large amount of literature has addressed the implications of least squares learning for popular dynamic stochastic general equilibrium (DSGE) models. Agents in a DSGE model that learn do not know the parameters of the model, and instead form expectations by collecting past data and compute least squares forecasts. In this paper I investigate statistical evidence for learning within the framework of a New Keynesian monetary model and examine the implications of incorporating learning on the predictions of the model

Recent papers have found that least squares learning can have important effects on output and inflation determination. Orphanides and Williams (2005b) use an estimated two equation monetary model and demonstrate with simulations of impulse response functions that least squares learning can lead to prolonged inflation following an inflation shock. Using the same model, Orphanides and Williams (2005a) find in another paper that learning on the part of monetary policy can possibly explain the period of stagflation during the 1970s. They suggest that the monetary authority was under-estimating the natural rate of unemployment during this time, and was therefore responding too aggressively to unemployment and not enough to inflation. They suggest that had the central bank responded to inflation instead of unemployment, lower inflation and unemployment would have resulted. Primiceri (2006) suggests that learning on the part of the central bank can explain both the run-up of inflation during the 1970s and the subsequent decline during the 1980s. He suggests that the monetary authority was under-estimating both the natural rate of unemployment and the degree of inflation persistence. Like Orphanides and Williams (2005a), he shows the resulting monetary policy leads to an increase in inflation, but as time progresses the central bank's expectations evolve. The central bank's expectations of the natural rate of unemployment and the degree of inflation persistence return their actual values and therefore the policy prescription becomes stabilizing, resulting in the moderation that occurred from the middle 1980s onward.

The results from these papers depend on a calibrated value for the constant learning gain, a parameter that is responsible for the speed in which expectations evolve, and therefore responsible for the impact learning can have on the dynamics of the model. Milani (2007) is the first paper to estimate the learning gain jointly with the parameters of a model. He finds an estimate for the learning gain which is very close to calibrated values that are popular in the literature. He estimates a standard three equation New Keynesian model and finds evidence in U.S. data that learning explains persistence in output and inflation better than habit formation and inflation indexation. Like the papers cited above, Milani makes specific assumptions about the initial conditions of agents expectations. Many of the initial conditions are set close to pre-sample ordinary least squares estimates. The exceptions are the degree of inflation persistence, which he assumes is equal to zero, and the sensitivity of output to inflation, which he assumes is higher than the pre-sample evidence.

The results of all of these studies depend on the assumptions for the initial conditions for agents and/or central bank's expectations. These initial conditions are sometimes backed by an economic justification or an argument that such a set of initial conditions accounts well for the data. In this paper, instead of suggesting a specific assumption for the initial expectations of agents, I examine a number of alternative methods for forming these initial conditions. These methods include using the rational expectations solution of the model and using least-squares estimation results from pre-sample data.

I extend the analysis of the existing empirical learning literature, and incorporate learning into a New Keynesian model with firm-specific capital and endogenous investment decisions, a model introduced by Woodford (2005). There are a number of motivations for extending the empirical analysis to the model with capital accumulation. Including capital in the model introduces data on another variable, aggregate investment, to be included in the estimation procedure. Secondly, introducing capital may alter how expectations are formed, since agents may use past data on capital to make their forecasts. Also, incorporating capital introduces more expectations into the model which may allow learning may play a bigger role. Finally, in a rich model with stochastic shocks to preferences, technology, and investment, one can determine the role learning has on the impact of such shocks, and the role the shocks play in explaining U.S. data.

Rational expectations and learning versions of the model are estimated by maximum likelihood. The findings of this paper indicate the learning gain is statistically significant which implies rejection of the null hypothesis of rational expectations and statistical evidence that expectations are adaptive. Examination of the other parameter estimates indicate that allowing for learning in the model can lead agents' consumption and investment decisions to be less dependent on expectations. Furthermore, estimated impulse response functions indicate that learning can lead to very different effects for the structural shocks depending on the information agents use for forming their forecasts and depending on the initial conditions for agents expectations at the beginning of the sample period. Despite these differences, examination of in-sample and out-of-sample forecast errors find that the learning models do not out-perform the rational expectations model in explaining the data.

The paper is organized as follows. Section 2 describes the details of the New Keynesian model with firm-specific capital. Section 3 describes the learning process and how learning is incorporated into the model. Section 4 describes the maximum likelihood procedure and the four cases for how initial conditions are constructed. Section 5 reports the results, and section 6 concludes.

2.2 Model

The New Keynesian model has been used extensively in monetary economics for analysis of theoretical and empirical issues and it is a convenient framework to examine the role of learning on output, consumption, investment, and inflation determination. Woodford (2003) provides a complete exposition of the model's micro-foundations, its many extensions, and implications for monetary policy. The model used in this paper is an extension of the standard three equation New Keynesian suggested by Woodford (2005) that incorporates endogenous investment decisions in a framework of firm-specific capital, where output is produced under constant returns using labor and firm-specific capital. Not only is firm-specific capital a more realistic assumption than a perfect rental market for capital, Woodford shows that allowing for firm-specific capital alters the coefficient on marginal cost in the Phillips curve in such a way that allows for greater price flexibility to be consistent with very small values of the coefficient, which is often seen in empirical work.

The model has a continuum of consumers types on the unit interval, and a continuum of intermediate goods producers on the unit interval, each producing a unique intermediate good. Each consumer type possesses a specific labor skill that can only be hired by a corresponding intermediate goods producer. It is assumed that there are many consumers in each consumer type so that consumers do not have market power over the wage. Production of intermediate goods also depend on capital goods which are firm-specific. Since a capital good in firm i cannot be used by another firm j, there is not a perfect capital rental market which would equalize the marginal product of capital across intermediate goods firms. Therefore each firm's labor demand and pricing decision will depend on its current capital stock, which in turn depends on the firm's entire past history.

All the intermediate goods are used to produce a single type of final good, but they are imperfect substitutes for each other in production; therefore intermediate goods producing firms are monopolistically competitive. Prices of intermediate goods are imperfectly flexible according to Calvo's (1983) pricing mechanism where a constant fraction of firms is able to re-optimize its price every period, and the firms selected to do so is randomly determined, independently of firms' histories or characteristics. This setup for sticky prices may seem unrealistic, but Roberts (1995) shows in a model without firm-specific capital that quadratic price adjustment cost, an alternative pricing friction suggested by Rotemberg (1982), yields the same solution as Calvo pricing. The same is not true with firm-specific capital. Under Calvo pricing, at any point in time, each firm will have a different pricing history and therefore a different capital stock. Each firm's relative capital stock will in turn affect the pricing decision. Under quadratic price adjustment costs, all firms face the same friction every period, and so all firms' price, labor, and investment decisions remain identical throughout time. Therefore, even though Calvo pricing may seem to be an unrealistic setting, it is a convenient framework to incorporate the realistic assumption of firm heterogeneity.

2.2.1 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,

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-\frac{1}{\sigma}} \xi_t \left(c_t - \eta c_{t-1}\right)^{1-\frac{1}{\sigma}} - \frac{1}{1+\mu} \mu_t n_t(i)^{1+\mu} \right], \tag{2.1}$$

subject to the budget constraint,

$$c_t + b_t(i) = \frac{1+r_{t-1}}{1+\pi_t} b_{t-1}(i) + \frac{w_t(i)}{p_t} n_t(i) + \Pi_t - \tau_t \tag{2.2}$$

where c_t, consumption at time t, is not indexed by individual type i since it is equal across all agents, ξ_t 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, μ_t is an aggregate labor supply shock, $b_t(i)$ 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, 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 pseudo intertemporal-elasticity of substitution,[1] $\eta \in [0, 1)$, which is the degree of habit formation, and $\mu \in (0, \infty)$ which is the inverse of the elasticity of labor supply. The appendix shows

[1] When there is no habit formation and labor supply is fixed, σ is exactly equal to the intertemporal elasticity of substitution.

that the first order conditions for the consumer lead to the log-linear Euler equation,

$$\hat{\lambda}_t = E_t\hat{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1}, \tag{2.3}$$

where a hat indicates the percentage deviation of the variable from its steady state.[2] Here, $\hat{\lambda}_t$ is the marginal utility of real income, given by,

$$\hat{\lambda}_t = \frac{1}{\sigma(1-\beta\eta)(1-\eta)}\left[\beta\eta E_t\hat{c}_{t+1} - (1+\beta\eta^2)\hat{c}_t + \eta\hat{c}_{t-1}\right] + \left(\hat{\xi}_t - \beta\eta E_t\hat{\xi}_{t+1}\right). \tag{2.4}$$

I assume that the preference shock, $\hat{\xi}_t$, follows the exogenous autoregressive process,

$$\hat{\xi}_t = \rho_\xi\hat{\xi}_{t-1} + \epsilon_{\xi,t}, \tag{2.5}$$

where $\epsilon_{\xi,t}$ is independently and identically with mean zero and variance given by σ_ξ^2.

When there is no habit formation, equations (2.3) and (2.4) lead to the standard IS equation,

$$\hat{c}_t = E_t\hat{c}_{t+1} - \sigma\left(\hat{r}_t - E_t\pi_{t+1}\right) + \hat{\xi}_t.$$

Habit formation is added to the model, because as equation (2.4) demonstrates, habit formation introduces a source of persistence that does not depend on learning. The larger is the degree of habit formation, the more current period marginal utility depends on past consumption. Since consumption is related to output in the market clearing condition, habit formation creates output persistence. Moreover, Fuhrer (2000) finds that habit formation leads to "hump shaped" impulse response functions, a phenomenon evident in the data.

2.2.2 Producers

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.

[2] A hat is omitted from inflation because, as demonstrated in the appendix, in order to derive the Phillips curve it is necessary to assume the steady state inflation rate is equal to zero.

The production function is given by,

$$y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di\right]^{\frac{\theta}{\theta-1}} \tag{2.6}$$

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,

$$y_t(i) = \left[\frac{p_t(i)}{p_t}\right]^{-\theta} y_t, \tag{2.7}$$

where $p_t(i)$ is the price of intermediate good i and p_t is the price of the final good. Substituting equation (2.7) into (2.6) leads to a consumption price index that holds in equilibrium,

$$p_t = \left[\int_0^1 p_t(i)^{1-\theta} di\right]^{\frac{1}{1-\theta}}. \tag{2.8}$$

Intermediate goods

The intermediate good is produced with labor and a unique type of capital good according to the constant returns to scale production function,

$$y_t(i) = z_t k_t(i)^{\alpha} n_t(i)^{1-\alpha} \tag{2.9}$$

where $k_t(i)$ is capital hired by firm i. For a given level of output, intermediate goods firms choose labor demand and rent capital to minimize real total cost,

$$C_t = \frac{w_t(i)}{p_t} n_t(i) + \rho_t(i) k_t(i), \tag{2.10}$$

where $\rho_t(i)$ is the rental price of capital good i. Log-linearizing the production function and summing over all intermediate goods firms leads to the log-linear aggregate production function,

$$\hat{y}_t = \hat{z}_t + \alpha \hat{k}_t + (1-\alpha)\hat{n}_t. \tag{2.11}$$

The appendix shows when firms hire optimal amounts of labor and capital, the average marginal cost among all the intermediate goods firms (in terms of the percentage deviation from the steady state) is given by,

$$\hat{s}_t = \frac{\alpha + \mu}{1 - \alpha}\hat{y}_t - \frac{\alpha(\mu + 1)}{1 - \alpha}\hat{k}_t - \hat{\lambda}_t - \frac{\mu + 1}{1 - \alpha}\hat{z}_t, \tag{2.12}$$

where $\hat{k}_t$ is the percentage deviation of the aggregate capital stock from its steady state. The technology shock is assumed to follow the exogenous stochastic process,

$$\hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t}, \tag{2.13}$$

where $\epsilon_{z,t}$ is independently and identically distributed with mean zero and variance given by σ_z^2.

Firm-specific capital goods

Capital goods firms maintain firm-specific capital stocks and rent the capital to the corresponding intermediate goods firm at a real price of $\rho_t(i)$ per unit of capital. This assumption in not essential and is purely used for notational convenience. This model supposes that the market for firm-specific capital is purely competitive, even though firm-specific capital cannot be sold to other firms. This assumption assures an optimal amount of investment in each firm-specific capital good which would be the same outcome if the intermediate goods firms were to invest and own the capital themselves instead of renting it.

Capital goods firms purchase the final good and convert it to a firm-specific capital good. The conversion from a final good to a firm-specific capital good is irreversible and is subject to a stochastic shock, ι_t, that is common to all capital goods. Let $I_t(i)$ denote the purchase of the final good for investment for capital good i, so that $\iota_t I_t(i)$ be the amount a purchase of $I_t(i)$ adds to the capital stock. The evolution of firm-specific capital i is given as,

$$k_{t+1}(i) = (1 - \delta)k_t(i) + \iota_t I_t(i) - \frac{\phi}{2}\left[\frac{k_{t+1}(i)}{k_t(i)} - 1\right]^2 k_t(i) \tag{2.14}$$

where $\delta \in (0, 1)$ is the capital depreciation rate and $\phi \in (0, \infty)$ is a capital adjustment cost parameter. When $\phi = 0$, there is no adjustment cost and capital net of depreciation increases by $\iota_t I_t(i)$. Log-linearizing equation (2.14) then integrating across all the firms leads to the following relationship between capital and investment:

$$\hat{k}_{t+1} = (1-\delta)\hat{k}_t + \delta\hat{I}_t + \delta\hat{\iota}_t \tag{2.15}$$

Capital goods firms choose investment to maximize the expected utility value of profits,

$$E_0 \sum_{t=0}^{\infty} \beta^t \lambda_t \left[\rho_t(i) k_t(i) - I_t(i)\right], \tag{2.16}$$

subject to equation (2.14). The appendix shows that profit maximization leads to the following evolution of the aggregate capital stock:

$$\begin{aligned}
\hat{\lambda}_t + \phi\left(\hat{k}_{t+1} - \hat{k}_t\right) = \beta(1-\delta)E_t\hat{\lambda}_{t+1} + \left(\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right)\left[(\mu+1)E_t\hat{y}_{t+1} - (1+\mu\alpha)\hat{k}_{t+1}\right] \\
+\beta\phi\left(E_t\hat{k}_{t+2} - \hat{k}_{t+1}\right) - \frac{(\mu+1)\left[1-\beta\left(1-\delta\right)\right]}{1-\alpha}E_t\hat{z}_{t+1} + \hat{\iota}_t - \beta(1-\delta)E_t\hat{\iota}_{t+1} \\
+\left[1-\beta(1-\delta)\right]E_t\hat{\mu}_{t+1},
\end{aligned} \tag{2.17}$$

The investment shock is assumed to follow the stochastic process,

$$\hat{\iota}_t = \rho_\iota \hat{\iota}_{t-1} + \epsilon_{\iota,t} \tag{2.18}$$

where $\epsilon_{\iota,t}$ is independently and identically distributed with mean zero and variance given by σ_ι^2.

Phillips Curve

The Phillips curve is a single equation that describes the relationship between inflation and output, as determined by the supply side of the economy when prices are sticky. The specific price friction employed in this paper is Calvo (1983) pricing. According to this method, only a random subset of intermediate goods firms are able to re-optimize their price in a given period. Allowing for inflation indexation, those 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 change their prices each period. Since the specific firms able to change their prices each period is randomly determined, ω^T is the probability a firm will not be able to change its price for T consecutive periods. A firm who is able to change its price maximizes the following present discounted utility value of profits earned while the firm is unable to change its price again:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_{t+T}(i)}{p_{t+T}} \right) y_{t+T}(i) - S\left[y_{t+T}(i) \right] \right\}, \tag{2.19}$$

where $S\left[y_{t+T}(i)\right]$ is the real total cost function of producing $y_{t+T}(i)$ units, given optimal decisions for labor and capital, and $p_{t+T}(i)$ is the firm's price in period $t+T$, given the firm has not yet been able to re-optimize its price. When there is a positive degree of inflation indexation, this price is determined by,

$$\log p_{t+T}(i) = \log p_{t+T-1}(i) + \gamma \pi_{t+T-1} \tag{2.20}$$

The appendix shows that the firms' optimal choices for prices in combination with equilibrium in the firm-specific capital goods market leads to the following Phillips curve,

$$\pi_t = \left(\frac{1}{1+\beta\gamma} \right) \left(\gamma \pi_{t-1} + \beta E_t \pi_{t+1} + \kappa \hat{s}_t \right) \tag{2.21}$$

where κ decreases as ω, the degree of price stickiness, increases. The parameter κ is also a function of other parameters of the model, but there is not a closed form expression for it. The appendix describes the full details of the derivation of the Phillips curve.

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,

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi E_t \pi_{t+1} + \psi_y E_t \hat{y}_{t+1}\right) + \epsilon_{r,t} \tag{2.22}$$

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 expected inflation, $\psi_y \in (0,\infty)$ is the feedback on the interest rate to expected output, 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. McCallum (1997), for example, argues that a policy rule that depends on current realizations of output and inflation is not operational. He suggests that using lagged realizations more accurately represent actual monetary policy since current quarter estimates for output and price levels are not available. Under rational expectations with full information, the policy rule above is also subject to this criticism. However, as will be seen in the next section, under learning expectations are formed by collecting past data, so the monetary policy rule above is operational.

2.2.3 Complete Model

The complete system has nine variables: consumption ($\hat{c}_t$), marginal utility of income ($\hat{\lambda}_t$), investment ($\hat{I}_t$), capital stock ($\hat{k}_t$), marginal cost ($\hat{s}_t$), output ($\hat{y}_t$), labor ($\hat{n}_t$), inflation (π_t), and the interest rate ($\hat{r}_t$). The demand side of the model consists of the Euler equation, (2.3), and the definition of the marginal utility of income, (2.4). The supply side of the model consists of the Phillips curve, (2.21), the definition of the marginal cost, (2.12), the evolution of capital, (2.17), the relationship between investment and capital, (2.15), and the production function (2.11). The model is completed with the monetary policy rule, (2.22), and the following log-linear goods market clearing condition,

$$\hat{y}_t = c_y \hat{c}_t + \delta k_y \hat{I}_t, \tag{2.23}$$

where c_y is the steady state consumption to output ratio and k_y is the steady state capital to output ratio. The appendix shows that k_y and c_y are given by,

$$k_y = \frac{\beta\alpha(\theta-1)}{\theta\left(1-\beta+\beta\delta\right)},$$

$$c_y = 1-\delta k_y.$$

There are five exogenous shocks in the model: the preference shock, ξ_t, whose evolution is given in equation (2.5); the technology shock, z_t, whose evolution is given in equation (2.13); the investment shock, ι_t, whose evolution is given in equation (2.18); the labor supply shock, μ_t, and the monetary policy shock, $\epsilon_{r,t}$.

2.3 Learning

The specific type of adaptive learning process considered in this paper is least squares learning. Under least squares learning, agents form expectations by collecting past data and computing least squares estimates. The specific type of least squares learning I use is constant gain learning, which is consistent with agents' forecasts based on weighted least squares, where more recent observations are given more weight, and the weights decline geometrically with the age of the observations. This is a popular assumption in the learning literature and is the same type of learning used by Orphanides and Williams (2005b) to explain inflation scares, Primiceri (2006) to explain the inflation volatility in the 1970s, and Milani (2007) to explain output and inflation persistence.

Constant gain least squares learning is arguably similar to how expectations are actually formed in the U.S economy. Least squares forecasts out-perform more complex economic models in out-of-sample forecasts, and the welfare of individuals who make output, consumption, and savings decisions depend on the accuracy of forecasts and not the ability to identify parameters of an econometric model, or the ability to make counter-factual predictions. These latter qualities, found in structural economic models, are desirable mostly by policy makers. The constant gain assumption can also be argued as realistic as it captures the idea that agents agents believe changes in the economy are possible, so that agents view more recent data as more likely to yield accurate forecasts than data from further in the past. I demonstrate in the next section that constant gain least squares is equivalent to a very specific type of weighted least squares which is not an actual popular estimation method. However, Evans and Honkapohja (2001) suggest that constant gain least squares is a good approximation for agents that use a "rolling window" of data. That is, agents do not use all the data as far back as possible, but form forecasts based on the most recent data for a given number of observations. This is very close to common practice, as empirical studies that forecast output and inflation typically use at most 50 years of data, despite annual data available from Johnston and Williamson (2007) for both these variables dating all they way back to the year 1790.

There is also a theoretical and empirical appeal to using constant gain learning. The theoretical appeal is that unlike with ordinary least squares, with weighted least squares the effects of learning persist in the long run. With ordinary least squares, as time progresses agents obtain more and more observations and so their sample sizes approach infinity. Therefore, the effect a single new observation has on the agents' estimation results disappears. Constant gain learning instead assumes that a new observation carries the same weight every period, regardless of how much time has progressed. The empirical appeal is that the degree to which learning affects the dynamics of the economy can be determined by estimating a single parameter, the learning gain. Moreover, with appropriate initial conditions for the learning process, constant gain learning nests the rational expectations framework, where rational expectations is the special case where the learning gain is equal to zero. Standard statistical tests that determine if a parameter is significantly different from zero can determine the statistical significance of learning, and formally reject or fail to reject the rational expectations hypothesis.

The log-linearized New Keynesian model in the previous section has the following general form:

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi v_t, \tag{2.24}$$

$$v_t = Av_{t-1} + \epsilon_t \tag{2.25}$$

where x_t is a vector of state variables (expressed as percentage deviations from their steady state), E_t^* refers to a possibly non-rational expectations operator, v_t is a vector of structural shocks, and ϵ_t is a vector of independently and identically distributed innovations to the shock process. In the New Keynesian model with firm-specific capital the state vector is given by, $x_t' = [\hat{y}_t\ \hat{c}_t\ \hat{k}_t\ \hat{\lambda}_t\ \hat{I}_t\ \hat{n}_t\ \hat{s}_t\ \pi_t\ \hat{r}_t]$, and the vector of shocks is given by, $v_t = [\hat{z}_t\ \hat{\iota}_t\ \hat{\xi}_t\ \hat{\mu}_t\ \epsilon_{r,t}]$

The solution of the model can be written as,

$$x_t = Gx_{t-1} + Hv_t. \tag{2.26}$$

To agents that learn with a correctly specified model, the actual values in the matrices G and H are unknown, but agents use the form of equation (2.26) to estimate future values of x_t by least squares. It is assumed that when agents begin period t, time t observations are not yet realized; therefore agents collect observations up through time period $t-1$. From this agents make least squares forecast, then make consumption, production, investment, and pricing decisions based on these expectations. Only after these decisions are implemented, that is at the end of time period t, do time t observations become available. This is both a realistic and mathematically simplifying assumption. The latest numbers from statistical agencies such as the Bureau of Labor Statistics are almost always at least one quarter old. It is of great mathematical convenience, because the term $E_t^* x_{t+1}$ in equation (2.24) is then only a function of observations through period $t-1$. Therefore, solving for x_t in terms of past state variables is straightforward. If instead $E_t^* x_{t+1}$ was a function of x_t, non-linear numerical methods would be needed to solve the model as least squares forecasts are non-linear.

To forecast x_{t+1}, agents estimate G and H by least squares using as regressors variables in the vector x_{t-1}, and the shocks included in v_t. Assuming agents have data available on shocks is not very realistic, but this assumption can be dropped. In Section 2.4 I estimate the models under both cases, so that when comparing the results from the learning and rational expectations models, it will be clear what results derive from the learning process, and what results derive from assuming that agents have a more limited information set.

Agents do not use all the variables in x_t as regressors, only those that correspond to non-zero columns in G. If an entire column in G is equal to zero, this implies that the past observation in the associated element in x_{t-1} does not influence x_t in the rational expectations solution. I assume agents know the structural form of the economy and therefore use as explanatory variables only the variables that have non-zero coefficients in G. In the New Keynesian model, the variables with non-zero coefficients in G include consumption, capital, the inflation rate, and the interest rate. The remaining variables in the model that are not used as explanatory variables are output, labor, marginal cost, marginal utility of income, and investment.

I assume agents also use a constant term in their least squares forecasts. The structural form of the model, (2.24), does not include a constant, but since this equation is written in terms of percentage deviations from the steady state, using a constant in agents' estimation equations implies agents do not know the steady state values of the economy.

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 estimate equation (2.26) by ordinary least squares, they form the estimate,

$$\Phi_t' = \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau X_\tau' \right)^{-1} \left(\frac{1}{t-1} \sum_{\tau=2}^{t} X_\tau Y_\tau' \right). \tag{2.27}$$

The ordinary least squares estimate Φ_t can be rewritten into the convenient recursive form:

$$\Phi_t = \Phi_{t-1} + g_t(Y_t - \Phi_{t-1}X_t)X_t' R_t^{-1}, \tag{2.28}$$

$$R_t = R_{t-1} + g_t(X_t X_t' - R_{t-1}), \tag{2.29}$$

where $g_t = 1/(t-1)$ is the learning gain.[3] 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. With ordinary least squares 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. Constant gain learning instead assumes that the learning gain g_t remains constant over time. This allows new observations to influence estimation results by the same weight throughout time. If the constant gain is equal to zero, the estimate Φ_t remains at its initial value throughout time. Given an initial value equal to the rational expectations solution, a zero constant learning gain implies rational expectations.

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. Agents' expectation of x_{t+1} is given by,

$$E_t^* x_{t+1} = \hat{g}_{0,t} + \hat{G}_t E_t^* x_t + \hat{H} E_t v_{t+1} \tag{2.30}$$

Note that equation (2.30) 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

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

squares. The time t estimate of x_t is given by,

$$E_t x_t^* = \hat{g}_{0,t} + \hat{G}_t x_{t-1} + \hat{H}_t v_t. \tag{2.31}$$

Plugging this into equation (2.30) yields,

$$E_t^* x_{t+1} = (I + \hat{G}_t)\hat{g}_{0,t} + \hat{G}_t^2 x_{t-1} + \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right) v_t. \tag{2.32}$$

Plugging the agents' forecast, (2.32), into the structural form of the model, (2.24), leads to the following actual law of motion for x_t:

$$x_t = \Omega_0^{-1}\Omega_2 \left(I + \hat{G}_t\right) \hat{g}_{0,t} + \Omega_0^{-1} \left(\Omega_1 + \Omega_2 \hat{G}_t^2\right) x_{t-1} + \Omega_0^{-1} \left[\Psi + \Omega_2 \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right)\right] v_t. \tag{2.33}$$

2.4 Estimation

2.4.1 Data

I estimate the model with quarterly U.S. data on real private consumption, real gross private domestic investment, consumer price index inflation, and the effective federal funds rate. Data is collected for 1970:Q1 through 2008:Q1 from the Federal Reserve Bank of St. Louis FRED database. Consumption and investment are put in per-capita terms by dividing the series by data on the civilian non-institutional population which is obtained from the Bureau of Labor Statistics.

In the New Keynesian model, consumption and investment are expressed in terms of the percentage deviation from their respective steady states. Since this data is non-stationary it is first de-trended by removing a common trend growth rate, similar to Ireland 2004a and 2004b. Even though productivity growth is not specified in the model, consumption and investment should have the same long run growth rate. This growth rate is determined by adding together consumption and investment, and taking the average growth rate over the sample. The average quarterly growth rate of output computed this way is $g_y = 0.0054$. De-trended consumption and investment is therefore determined according to,

$$CONS_t^* = \frac{CONS_t}{(1+g_y)^t},\ INV_t^* = \frac{INV_t}{(1+g_y)^t},$$

where $CONS_t$ and INV_t are the raw data on consumption and investment, and $CONS_t^*$ and INV_t^* denote the data with the trend growth rate removed.

2.4.2 Initial conditions

I estimate the model under four different cases for how expectations are formed. Case 1 is rational expectations, and the other cases are learning with different assumptions for initial expectations and what explanatory variables agents use to make their least squares forecasts. Case 2 can be viewed as the closest to rational expectations. Agents learn according to constant gain least squares, but the initial values for the learning matrices Φ and R are equal to the rational expectations solution. Furthermore, agents have the same information as agents with rational expectations, which means they include realizations of structural shocks among the other explanatory variables. When the constant learning gain is equal to zero, Case 2 is equivalent to Case 1.

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.

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. This is similar to how Milani (2007) initializes the learning matrices, but he uses estimates from a first order vector autoregression using ordinary least squares, which is consistent instead with a decreasing learning gain. In this paper, the initial conditions for the learning process are consistent with the constant learning gain which is estimated jointly with the other parameters of the model.

Equations (2.28) and (2.29) 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 coefficient matrix is given by,

$$\Phi_t = \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \right)^{-1} \left(\sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} Y'_{t-\tau} \right) \tag{2.34}$$

In the New Keynesian model with firm-specific capital, $Y'_t = [\hat{c}_{t-1} \ \hat{k}_{t-1} \ \pi_{t-1} \ \hat{r}_{t-1}]$, and $X'_t = [1 \ \hat{c}_{t-2} \ \hat{k}_{t-2} \ \pi_{t-2} \ \hat{r}_{t-2}]$. In this specification of the model, some of the data agents use are not directly observed by the econometrician. Consumption is expressed as percentage deviations from the steady state, and capital stock is not directly observable. For a given estimate of the consumption to output ratio and the steady state level of output, pre-sample data on aggregate consumption is put in terms of the percentage deviation from the steady state according to,

$$\hat{c}_t = \frac{CONS_t^* - c_y y^*}{c_y y^*}, \tag{2.35}$$

where c_y is the steady consumption to output ratio, one of the New Keynesian model parameters to be estimated, and y^* is the steady state level of output which will be calibrated, as discussed in the next subsection.

The inflation rate is directly observable by the econometrician, but to make solving the New Keynesian model tractable, it was assumed in Section 2 that the steady state inflation rate is equal to zero. Since this is unlikely to be the case in the data, let π^* denote the annualized steady state inflation rate, expressed as a percentage. Let INF_t denote the annualized quarterly inflation rate measured from CPI data. This data is mapped to pre-sample data for π_t according to,

$$\pi_t = \frac{1}{400} \left(INF_t - \pi^* \right). \tag{2.36}$$

The steady state inflation rate, π^*, will also be calibrated as discussed in the next section.

The interest rate in the model is also expressed as a deviation from its steady state. The steady state real gross interest rate in the New Keynesian model is given by β^{-1}. Let r^* denote the annualized quarterly steady state interest rate so that $r^* = 400(\beta^{-1} - 1)$. Let FF_t denote data on the annualized quarterly federal funds Rate. Pre-sample data on the federal funds rate can therefore be transformed to pre-sample data for agents according to

$$\hat{r}_t = \frac{1}{400}\left(FF_t - r^* - \pi^*\right). \tag{2.37}$$

Data for the U.S. capital stock is difficult to measure, but using the New Keynesian model, data for percentage deviation of capital from its steady state level can be composed from data on the deviation of investment from its steady state level. Recall equation (2.15) describes the evolution of the capital stock in terms of the percentage deviation from the steady state:

$$\hat{k}_{t+1} = (1-\delta)\hat{k}_t + \delta\hat{I}_t + \delta\hat{\iota}_t. \tag{2.38}$$

For a given initial value for $\hat{k}_t$ in the pre-sample, a simulated path for the investment shock, $\hat{\iota}_t$, and pre-sample data on the percentage deviation of investment from the steady state, $\hat{I}_t$, a pre-sample series of $\hat{k}_t$ can be constructed. Investment in the model is expressed as a percentage deviation from its steady state. Similar to consumption data, pre-sample data on gross private domestic investment can be transformed to pre-sample data for $\hat{I}_t$ according to,

$$\hat{I}_t = \frac{INV_t^* - (1-c_y)y^*}{(1-c_y)y^*}, \tag{2.39}$$

where $(1-c_y)y^*$ is the steady state level of investment.

I suppose at the beginning of the pre-sample capital is equal to its steady state value so that the pre-sample initial value for $\hat{k}_t$ is set equal to zero. The investment shock series is generated equation (2.18) given the variance of the investment shock, σ_ι^2, which is estimated along with the rest of the parameters of the model in the maximum likelihood procedure.

Pre-sample quarterly data is collected for 1954:Q3 through 1969:Q4 and is transformed to pre-sample data for $\hat{c}_t$, $\hat{k}_t$, π_t, and $\hat{r}_t$ according to equations (2.35), (2.36), (2.37), (2.38), (2.18), and (2.39). The initial condition for Φ_0 is then computed using equation for the weighted least squares procedure, equation (2.34).

2.4.3 Maximum Likelihood Procedure

I estimate the model by maximum likelihood following the Kalman filter procedure outlined in chapter 13 of Hamilton (1994). This procedure involves rewriting the model into state space form. The state equation is a linear equation describing the entire New Keynesian model including the learning mechanism. The equations governing the state are the actual law of motion for x_t, given in equation (2.33), and the evolution of the structural shocks given in equation (2.25). Equation (2.33) can be rewritten more compactly as,

$$x_t = b_t + F_t x_{t-1} + M_t v_t, \tag{2.40}$$

where vector b_t and matrices F_t and M_t are given by,

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

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

$$M_t = \Omega_0^{-1} \left[\Psi + \Omega_2 \left(\hat{G}_t \hat{H}_t + \hat{H}_t A\right)\right]$$

This equation can be combined with equation (2.25) into the single state equation,

$$x_t^* = b_t^* + F_t^* x_{t-1}^* + \epsilon_t^*, \tag{2.41}$$

where $x_t^* = [x_t' \; v_t']'$ and,

$$F_t^* = \begin{bmatrix} F_t & M_t A \\ 0 & A \end{bmatrix},$$

$$b_t^* = \begin{bmatrix} b_t \\ 0 \end{bmatrix},$$

$$\epsilon_t^* = \begin{bmatrix} M_t \epsilon_t \\ \epsilon_t \end{bmatrix}.$$

The variance of ϵ_t^* is given by,

$$Var(\epsilon_t^*) = \begin{bmatrix} M_t \Sigma M_t' & M_t \Sigma \\ \Sigma M_t' & \Sigma \end{bmatrix},$$

where Σ is a diagonal matrix with the variance of the structural shocks along the diagonal.

The observations equations are given by,

$$\begin{aligned} CONS_t^* &= c_y y^* + c_y y^* \hat{c}_t \\ INV_t^* &= (1 - c_y) y^* + (1 - c_y) y^* \hat{I}_t \\ INF_t &= \pi^* + 400 \pi_t \\ FF_t &= \pi^* + 400 \left(r^n + \hat{r}_t \right), \end{aligned} \tag{2.42}$$

The likelihood is maximized with respect to the following vector of parameters,

$$\Theta_2 = [\eta \ \sigma \ \mu \ c_y \ \phi \ \gamma \ \rho_r \ \psi_y \ \psi_\pi \ \rho_z \ \rho_\iota \ \rho_\xi \ \sigma_z \ \sigma_\iota \ \sigma_\xi \ \sigma_r \ g].$$

Several parameters are calibrated instead of estimated. The discount factor, β, is set equal to 0.9925 which corresponds to a steady state real interest rate approximately equal to 3%. The depreciation rate, δ, is set equal to 0.025 which corresponds to an approximate annual depreciation rate of 10%.

The steady state level of inflation, π^*, is set equal to 3.67, the average inflation rate over the entire pre-sample and sample period. The steady state level of output y^* is set equal to the average of $CONS_t^* + INV_t^*$ over the sample and pre-sample period, which is computed to be $y^* = 14,355$. This is the average de-trended estimate for real per-capita output, when considering only investment and consumption and ignoring other spending that contributes to real GDP such as government spending and net exports.

Preliminary estimation results led to unreasonably low values for κ, the slope on the Phillips curve that depends on the degree of price flexibility, and α, the capital-share of income. Ireland (2004b) also reports difficulty in obtaining sensible estimates for the Phillips curve slope using maximum likelihood, calibrates this parameter to $\kappa = 0.1$. Smets and Wouters (2005) in very rich New Keynesian with capital accumulation report difficulty in estimating the capital-share of income and calibrate it to $\alpha = 0.24$. I follow each of these papers and use the calibrations.

Finally, data on total hours of labor is not used in the estimation procedure, so the labor supply shock $\hat{\mu}_t$ is suppressed, which leaves ρ_μ and σ_μ out of the estimation.

2.5 Results

2.5.1 Parameter Estimates

The parameter estimates for the four cases for expectations formation are given in Table 2.1. In this subsection I look at each case in turn.

Case 1: Rational Expectations

The first two columns of Table 2.1 show the parameter results for rational expectations. The parameter estimates for the sources of persistence are markedly low. The estimate for degree of habit formation is $\eta = 0.1060$, the estimate for the degree of price indexation is $\gamma = 0.3624$, and the estimate of the degree of monetary policy persistence is $\rho_r = 0.1945$. These are quite small compared to much of the empirical literature for dynamic macroeconomic models. For example, for U.S. data, Smets and Wouters (2005) find a degree of habit formation equal to 0.69 and a degree of price indexation equal to 0.66. Milani (2007) also predicts significant degrees for these sources of persistence, but only when estimating his model under rational expectations.

Some empirical work finds similar evidence for weak sources of persistence. Ireland (2004b) estimates by maximum likelihood a three equation New Keynesian model that is augmented with sources of persistence for output and inflation and finds estimates of these parameters statistically insignificantly different from zero. Cogley and Sbordone (2005) estimate the Phillips curve side of the model and find a median estimate for the degree of inflation indexation equal to zero. Despite the weak evidence for persistence from habit formation and price indexation, the persistence in the structural shocks in the model are all very significant. All these estimates are in excess of 0.9, and strongly significantly different from zero.

The estimated intertemporal elasticity of substitution is approximately equal to $\sigma = 0.1603$, which is rather small compared to other findings. For example, Smets and Wouters (2005) estimate the inverse elasticity equal to 1.62, implying the elasticity is approximately 0.61. Giannoni and Woodford (2003) find a similar estimate of 0.66. The estimate of this parameter may depend crucially on the assumed expectations mechanism since it measures the response of current period consumption decisions to changes in the expected real interest rate. This matter is further examined with Cases 2, 3 and 4 below.

The point estimate for the inverse elasticity of labor supply is approximately $\mu = 30.67$ which implies labor supply is very inelastic. One place in the model this parameter enters is in the evolution of capital stock implied by the optimal investment decision, given in equation (2.17). This equation illustrates that large values for μ imply relatively small responses in k_{t+1}, and therefore current period's optimal choice for investment, to changes in expectations for future output and future technology shocks. Due to the relationship of this parameter and expectations, the estimate of this parameter is also possibly dependent on the assumed expectations mechanism.

The estimate for the cost of capital adjustment is equal to 7.68. This is relatively close to the finding in Smets and Wouters (2005) and the calibration used in Woodford (2005). Dividing every term in equation (2.17) illustrates that this parameter also measures how responsive investment decisions are to changes in expectations. The larger is the cost of adjusting capital, the less responsive is investment to changes in expected output, expected technology shocks, and expected investment shocks.

Finally the monetary policy parameters indicate a strong response of the Federal Funds rate to expected inflation, $\psi_\pi = 2.1212$, and virtually no response to the deviation of expected output from its steady state, $\psi_y = 0.0000$. The more than one-to-one response of the Federal Funds rate to inflation is a common finding in the literature using data during the Volcker-Greenspan period after 1982. Lubik and Schorfheide (2004) find evidence that the response was smaller before this period, but since the sample period for this paper begins in 1970, most of the sample is data from a period where U.S. monetary policy is widely viewed as aggressively targeting inflation to promote macroeconomic stability. The absent response to output is somewhat different than what is found in the literature. Smets and Wouters (2005) use a more rich specification for the Taylor rule and find a similar weak response to current output, but a stronger response to lagged output.

Case 2: Learning with RE Initial Conditions

The next two columns of Table 2.1 show the results for learning when agents expectations at the beginning of the sample are set equal to the rational expectations solution, and agents use data on all the shocks to form their expectations, which is the same information set for rational expectations. Rational expectations is the special case when the learning gain, g, is equal to zero.

The point estimate for the learning gain for Case 2 is 0.0240 and is statistically significantly different from zero, which implies significant statistical evidence to reject the null hypothesis that expectations are rational. This is close to commonly calibrated values in the learning literature, for example Primiceri (2006) uses a calibration equal to 0.015 and Orphanides and Williams (2005b) uses a value of 0.05. This is also very close to Milani's (2007) estimate of 0.018.

While constant gain learning implies that agents use weighted least squares with a possibly indefinitely large sample, Evans and Honkapohja (2001) and Sargent (1999) among others have suggested that constant gain learning should closely resemble learning with ordinary least squares in which agents use a rolling window of approximately $1/g$ observations, since the additional weight given to a new observation in such a framework is also constant and equal to g. With this interpretation of the learning gain, the parameter estimate implies that agents look at data from the last 41.6 quarters, or almost 10.5 years.

Three parameter estimates, those for the intertemporal elasticity of substitution (σ), the inverse elasticity of labor supply (μ), and the cost of capital adjustment (ϕ), are very different from Case 1. All these parameters move in the direction that causes consumption and investment decisions to be less responsive to changes in expectations. The estimate for the intertemporal elasticity of substitution is approximately $\sigma = 0.0513$ which is about one third the size of the estimate under rational expectations. This implies that learning leads to the prediction that consumption decisions are less responsive to changes in the expected real interest rate.

The estimate for the inverse elasticity of labor supply is $\mu = 0.0499$, markedly different from the estimate under rational expectations. This implies a very elastic labor supply, and therefore a much smaller response in investment decisions to changes in expected future output and technology shocks. The direction for this different finding is intuitive. Given the finding that learning models predict agents decisions are less responsive to changes in expectations, changes in investment decisions will be less responsive if labor supply is elastic, since then firms can respond by changing their hiring decisions with relatively small changes in the real wage.

The estimated cost of capital adjustment is approximately $\phi = 24.8826$, much higher than predicted by rational expectations. Since adjusting capital is relatively more expensive, investment decisions are less responsive to changes in expectations about output, technology shocks, and investment shocks.

Finally, this learning framework implies a lesser monetary policy response to expected inflation, and a larger response to expected output. The finding that monetary policy is more responsive to output in the learning framework is related to Smets and Wouters (2005) finding that monetary policy responds more to lagged output than concurrent output. McCallum (1997) also argues that it is more realistic to suppose the monetary authority only has information on lagged output. Under learning, agents collect information up through the previous period, therefore expectations of future output depend completely on past data of output, not on current realizations.

Case 3: Learning with RE Initial Conditions and Unobservable Shocks

The next two columns of Table 2.1 show the estimation results when agents learn, but with a more limited information set than Cases 1 and 2. In Case 3 agents do not collect data on past realizations of structural shocks to use for explanatory variables in their forecasts. For the first period in the sample, the coefficients on the remaining explanatory variables are initialized to the rational expectations solution.

The estimate for the learning gain is approximately 0.0236 and is again statistically significantly different from zero. This does not imply a formal rejection of rational expectations as before, since the rational expectations framework is no longer a special case, but it does imply statistical evidence that expectations are adaptive. The estimate for the learning gain implies that agents use approximately 42.4 past quarters of data to form their expectations, or about 10.5 years.

Some notable differences in the estimates from the previous expectations frameworks again include the intertemporal elasticity of substitution, the inverse elasticity of labor supply, and cost of adjusting capital. The intertemporal elasticity of substitution is even smaller than the previous two cases, which means this framework for learning predicts an even smaller response of consumption decisions to changes in the expected real interest rate.

The estimate for the inverse of labor supply is $\mu = 2.0877$ which implies labor is inelastic. This is still far below the estimate under rational expectations, but much larger than the estimate under the learning framework in case 2. Therefore when not supposing that agents observe structural shocks to form their expectations, investment decisions are more sensitive to changes in expected future output relative to when agents do collect data on structural shocks. When agents only use past data on observable macroeconomic variables, current period shocks have no influence on expectations, so agents forecasts should be less volatile. Given a less volatile series for expectations, it is expected that an equivalent change in expectations should have a larger impact on agents' decisions than with more volatile expectations that is predicted when agents do observe shocks. Such an explanation is consistent with the larger estimate for μ, but possibly contradicts the finding that the estimate for σ is smaller. The estimate for the standard deviation of the preference shock is much higher in Case 3 than Cases 1 and 2. This causes consumption decisions to be more volatile which likely influences the estimate for the intertemporal elasticity of substitution.

The estimate for the cost of adjusting capital is approximately $\phi = 26.83$, which is not significantly different than under case 2, but still significantly greater than the estimate under rational expectations.

Case 4: Learning with Pre-Sample Initial Conditions

The final learning framework assumes agents have the same information set as Case 3, but initial conditions for the learning matrices are set equal to weighted least squares estimates obtained from pre-sample data. The estimate for the learning gain is approximately $g = 0.0381$, which is slightly larger than the previous estimates. This again suggests statistical evidence than expectations are adaptive. This estimate implies that agents use approximately 26.25 quarters, or about 6.5 years, of past data to form their forecasts.

The estimate for the elasticity of substitution is approximately $\sigma = 0.1220$ which is not very different than predicted under rational expectations. However, the estimate for elasticity of labor supply and cost of adjusting capital tell a similar story as Case 2. Relative to rational expectations, this framework for learning implies investment decisions are much less responsive to changes in expectations for future output.

2.5.2 Performance Comparison

The top part of Table 2.2 reports the in-sample root mean squared error (RMSE) of the residuals, that is the one-period ahead forecast errors for consumption, investment, inflation, and the federal funds rate. The table shows no clear improvement in performance by including learning in the model. The expectations framework that best described consumption and inflation is the learning model with observable shocks. The framework that best describes investment is the learning model with unobservable shocks. The framework that best describes the federal funds rate is rational expectations, yet the RMSE for the federal funds rate is very close across the first three cases. Case 4, learning with initial expectations based on pre-sample data, is the worst performing model for consumption, investment, and the federal funds rate; and is the second worst performing model for the inflation rate.

The bottom part of Table 2.2 shows the first-order autocorrelation for the squared residuals. If the New Keynesian model is correctly specified, and accurately captures the run-up of macroeconomic volatility in the 1970s and the subsequent decline after 1982, one would expect no autocorrelation in the residuals. Cases 1, 2, and 3 show autocorrelation insignificantly different from zero for the volatility in consumption and investment residuals. This suggests the New Keynesian model with firm-specific capital can adequately explain the dynamics of these variables under learning or rational expectations. However, in Cases 1, 2, and 3, the autocorrelation for the volatility in the inflation rate and federal funds rate residuals are significantly positive, which implies the models do not accurately capture the changing volatility in the data for inflation or monetary policy.

These results flip for Case 4. Here there is insignificant autocorrelation in the volatility for inflation and the federal funds rate residuals, but a higher degree of autocorrelation for the volatility of consumption and investment residuals.

To better understand the relative performance of each expectations framework over the sample period, Figure 2.1 shows the plots the each series of forecast errors. Periods of recession in U.S. history are shaded. The number in parentheses is the correlation of the series of forecast errors of the respective learning model with the series predicted under rational expectations. The forecast errors for all four variables for Cases 2 and 3 are highly correlated with rational expectations. The largest forecast errors are made during the 1970s and early 1980s, a period many agree was marked by excessive macroeconomic volatility. The largest forecast errors for the federal funds rate are made just after 1980 when Paul Volcker became chairman of the Federal

Reserve and began to aggressively counter-act very high levels of inflation.

The forecast errors under Case 4 are significantly less correlated with the rational expectations case, but it tells the same qualitative story. The model makes the largest forecast errors for consumption, investment, and inflation during the 1970s and early 1980s.

To get a further understanding of the relative performance of each model I next examine how they compare in out-of-sample extended forecasts. To do this I first estimate the model using the sub-sample 1970:Q1 through 1989:Q4, then use this set of parameters to make out-of-sample extended forecasts for forecast horizons 1 period ahead through 12 periods ahead for the remainder of the sample, 1990:Q1 through 2008:Q1. Figure 2.2 shows the RMSE's for consumption, investment, inflation and the federal funds rate for each of the four expectations frameworks. The horizontal axis is the forecast horizon and the vertical axis is the root mean squared error. The vertical axis is logarithmic in order for the RMSE's for each expectations framework to show nicely on each graph.

The results from Figure 2.2 show the rational expectations model performs nearly as well out-of-sample as most of the learning models. The worst performing model for investment, inflation, and the federal funds rate is Case 2, the learning model arguably closest to rational expectations. Recall, in this framework agents include data on structural shocks to form their forecasts, and expectations at the beginning of the sample are set equal to the rational expectations solution. In this framework the learning matrices, Φ_t and R_t evolve with realizations of the shock processes. Under rational expectations, agents expectations also depend on structural shocks, but the learning matrices are constant throughout time, equal to the rational expectations solution. Out-of-sample forecasts for the structural shocks generate imprecise forecasts for the learning matrices, and therefore poor out-of-sample forecasts. This result would not be expected for a learning gain very close to zero. If the learning gain were very small, the learning gain matrices would be very slow to evolve.

2.5.3 Structural Shocks

Despite the similar performance of the four models for in-sample and out-of-sample forecast errors, the choice for how expectations are formed makes a difference for predictions for the structural shocks. Figures 2.3, 2.4, 2.5, and 2.6 show the impulse responses to consumption, investment, inflation, and the interest rate for each shock under each of the expectations frameworks.

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 2.3, 2.4, 2.5, and 2.6 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.

The figures show persistent impacts on consumption, investment, and the interest rate from preference, technology, and investment shocks under rational expectations. In Case 2 the shocks cause large, long lasting, oscillatory effects on investment, inflation, and the interest rate. This oscillatory behavior is likely what causes the large out-of-sample forecast errors for Case 2. The figures show when expectations are initialized based on pre-sample data (Case 4), the responses to the four shocks also cycle back and forth but are much more jagged, more volatile, but shorter lasting.

When agents do have data on structural shocks, the effects of structural shocks can have opposite impacts as when agents do observe structural shocks. For example, the first and second rows of Figure 2.3 show when there is a positive preference shock and agents can observe the shock, and therefore recognize that it is temporary, the interest rate increases in response to the increase in consumption and investment decreases. In Case 3 where agents learn and do not observe the shock, the increase in consumption demand causes agents to expect permanently higher consumption and therefore increase investment.

Figure 2.5 shows when there is a positive investment shock, in Case 2 there is an increase in investment demand as agents expect investment to be more productive. In Case 3, agents cannot view the shock and so do not initially expect investment to be more productive. The shock does cause an unexpected increase in the capital stock which decreases the marginal product of capital and therefore decreases the demand for investment.

Figure 2.6 shows the impact of a contractionary monetary policy shock. In Cases 1 and 2 when agents observe the monetary policy shock, the demand for consumption and investment decrease. When agents do not observe the shock, the stochastic behavior of the interest rate causes agents to alter their expectations for the behavior of monetary policy. The supply for output increases causing an increase in consumption and investment and a decrease in inflation. The negative response to inflation then causes the monetary authority to actually decrease the interest rate after the initial positive shock.

To understand how these impulse responses influence the predictions of the model, Figure 2.7 shows the smoothed estimates[4] for the evolution of the structural shocks over the sample period. Again, periods of U.S. recession are shaded and the number in parentheses is the correlation of each shock with the prediction under rational expectations. The evolution of the preference shock under Cases 1, 2, and 3 are all very similar. In each of these cases the preference shock is near is lowest during the 1980 and 1981 recessions. Other than this instance, the preference shock seems to have little correlation with onset of recessions. The same is not true for Case 4. The preference shock under Case 4 is completely uncorrelated with the prediction under rational expectations, and shows only small dips during recessionary periods.

The evolution of the technology shocks in Cases 3 and 4 are highly correlated with the rational expectations prediction, but the Case 2 evolution of technology shocks is completely uncorrelated. Recall from Figure 2.4 that in Case 2, technology shocks cause long-lived oscillatory behavior in consumption, investment, and inflation,

[4] Kalman filter smoothing is computed using the techniques described in de Jong (1989).

whereas under rational expectations and Case 3, the responses of these variables are in one direction and long lived. Despite the lack of correlation in the technology shock in Case 2, all cases have a similar qualitative story. All recessionary periods are marked by low realization of the technology shock. The difference in Case 2 is that the drop in the technology shock actually precedes the recessions, instead of dropping during the recessions as in other cases. In fact from 1980 through 1982 the technology shock is actually recovering in Case 2, but remaining low in all the other frameworks.

The evolution of the investment shock is very different among the four cases. Under rational expectations, the recessions in the early 1980s are characterized by large positive investment shocks. Recall from Figure 2.5 that a positive investment shock under rational expectations causes a negative response to consumption and a positive response to investment, inflation, and the interest rate. These responses are consistent with the United States experience during the late 1970s and early 1980s, a period marked by high unemployment and high inflation, and a time when the Federal Reserve began to drastically increase the federal funds rate in response to high inflation. The same is not true for Case 2. Figure 2.5 shows that consumption responds positively to an investment shock, the initial positive response to investment is much larger, and the responses to investment and inflation are shorter lived. The evolution of the investment shock therefore looks somewhat different over the sample period. The investment shock still reaches peaks during this time, but it is not as dramatic as under rational expectations.

The evolution of the investment shock in Cases 3 and 4 are similar to each other, but very different from Cases 1 and 2. Here the recessions in the early 1980s are characterized by negative investment shocks. When agents do not use structural shocks to form their expectations, Figure 2.5 shows that positive investment shocks lead to a drop in inflation, a sustained positive response to consumption, and a fall in the interest rate. The early 1980s were instead characterized by decreases in consumption and increases in interest rates, therefore, the investment shock falls during this period, which is the opposite of the finding in Cases 1 and 2.

2.6 Conclusion

Constant gain learning is found to not significantly out-perform rational expectations in the context of a standard Keynesian model that is extended to account for endogenous firm-specific capital accumulation. Three different learning frameworks are examined that differ based on the initial conditions for agents' expectations and the information set available to agents when making their forecasts. The models are estimated by maximum likelihood and the results indicate that learning provides minimal to no improvement in the fit of the model to U.S. data. When the learning procedure is initialized using least squares estimation results from pre-sample data, the model actually performs the worst as measured by in-sample forecast errors. Out-of-sample extended forecast errors show the worse performing model is the learning framework in which agents do have information on all structural shocks and expectations at the beginning of the sample are set equal to rational expectations. Despite the mixed results for the fit for the various models to the data, this paper does find a constant learning gain statistically significantly different from zero in all cases, which implies statistical evidence that expectations are not rational and are adaptive.

Impulse response functions and smoothed estimates of the structural shocks reveal that each expectations framework provides very different explanations for the impacts a structural shock can have and therefore each model predicts different evolutions for the structural shocks over the sample period. The impulse response functions show when agents learn and have data on structural shocks, there can be long-lived oscillatory effects on consumption, investment, and inflation. When agents do not observe the structural shocks, some of the impulse responses are reversed. These different responses to structural shocks lead to different evolutions for the structural shocks over the sample period, especially for the technology and investment shocks.

References

Bullard, J., and J. Duffy (2004): "Learning and Structural Change in Macroeconomic Data," Federal Reserve Bank of St. Louis Working Paper 2004-016A.

Bullard, J., and S. Eusepi (2005): "Did the Great Inflation Occur Despite Poicymaker Commitment to a Taylor Rule?," Review of Economic Dynamics, 8, 324–359.

Bullard, J., G. W. Evans, and S. Honkapohja (2005): "Near Rational Exuberance," Federal Reserve Bank of St. Louis Working Paper 2004-025B.

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

Carceles-Poveda, A., and C. Giannitsarou (2005): "Adaptive Learning in Practice," Working Paper.

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.

——— (2008): "Expectations, Learning and Monetary Policy: An Overview of Recent Research," Centre for Dynamic Macroeconomic Analysis Working Paper CDMA08/02.

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. (2004a): "A method for taking models to the data," Journal of Economic Dynamics and Control, 28, 1205–1226.

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

——— (2005): "Irrational Expectations and Econometric Practice," Federal Reserve Bank of Atlanta Working Paper 2003-22.

——— (2006): "Changes in the Federal Reserve's Inflation Target: Causes and Consequences," Working Paper.

Johnston, L. D., and S. H. Williamson (2007): "The Annual Real and Nominal GDP for the United States, 1790 - present," URL: http://eh.net/hmit/gdp/.

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 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.

Mark, N. C. (2005): "Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dyanamics," NBER Working Paper 11061.

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

Milani, F. (2005): "Learning, Monetary Policy Rules, and Macroeconomic Stability," Working Paper.

——— (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.

Roberts, J. M. (1995): "New Keynesian Economics and the Phillips Curve," Journal of Money, Credit and Banking, 27, 975–984.

Rotemberg, J. (1982): "Sticky Prices in the United States," Journal of Political Economy, 90, 1187–1211.

Rotemberg, J., and M. Woodford (1997): "An Optimization Based Econometric Framework for the Evaluation of Monetary Policy," In: Bernanke, B.S. and J. Rotemberg (Eds.), NBER Macroeconomics Annual. MIT Press.

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

Sims, C. (2000): "Solving Linear Rational Expectations Models," Unpublished manuscript.

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.

Smets, F., and R. Wouters (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.

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

——— (2005): "Firm Specific Capital and the New Keynesian Phillips Curve," International Journal of Central Banking, 1, 1–46.

Table 2.1: Maximum Likelihood Parameter Estimates: Sample 1970: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	0.1060	0.0272	0.1289	0.0399	0.1224	0.0264	0.2728	0.0232
Intertemporal Substitution	σ	0.1603	0.0342	0.0513	0.0072	0.0157	0.0001	0.1220	0.0175
Inverse Elas. Labor Supply	μ	30.6713	8.9559	0.0499	0.0771	2.0877	0.4286	0.3324	0.2817
Cons. / Output Ratio	c_y	0.8339	0.0038	0.8672	0.0024	0.8361	0.0000	0.8289	0.0028
Capital Adjustment Friction	ϕ	7.6832	1.4003	24.8826	1.4036	26.8332	4.1201	26.9755	1.7912
Price Indexation	γ	0.3624	0.1929	0.0000	0.0349	0.5236	0.0852	0.6090	0.1214
MP Persistence	ρ_r	0.1945	0.0497	0.6592	0.0141	0.7250	0.0280	0.0956	0.1061
MP Output	ψ_y	0.0000	0.0172	0.0576	0.0090	0.0458	0.0092	0.0041	0.0131
MP Inflation	ψ_π	2.1212	0.1893	1.4448	0.0658	1.7735	0.1278	1.8491	0.0974
Preference Shock Pers.	ρ_ξ	0.9826	0.0069	0.9925	0.0040	0.9636	0.0070	1.0000	0.0000
Technology Shock Pers.	ρ_z	0.9668	0.0058	0.6741	0.0254	0.9638	0.0193	0.9506	0.0121
Investment Shock Pers.	ρ_ι	0.9060	0.0151	0.9301	0.0084	0.9297	0.0103	0.9234	0.0173
Pref. Shock Std. Dev.	σ_ξ	0.0926	0.0252	0.1647	0.0211	0.6165	0.0511	0.6879	0.0446
Tech. Shock Std. Dev.	σ_z	0.0104	0.0003	0.0199	0.0018	0.0294	0.0037	0.0730	0.0191
Inv. Shock Std. Dev.	σ_ι	0.0206	0.0021	0.1130	0.0149	0.0387	0.0016	0.0683	0.0069
Policy Shock Std. Dev.	σ_r	0.0027	0.0003	0.0032	0.0001	0.0036	0.0001	0.0053	0.0003
Learning Gain	g	–	–	0.0240	0.0043	0.0236	0.0026	0.0381	0.0038

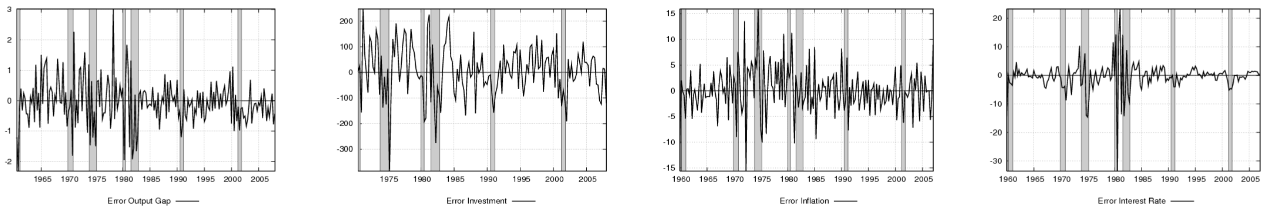
Table 2.2: Model Fit Comparisons

	Root Mean Squared Error			
	Case 1	Case 2	Case 3	Case 4
Consumption	110.6570	87.6595	106.8923	143.3313
Investment	99.9744	105.0477	87.0809	132.0726
Inflation	2.6073	2.4222	3.1500	2.9596
Federal Funds Rate	1.3661	1.3930	1.3714	1.9271
	Autocorrelation Squared Error			
	Case 1	Case 2	Case 3	Case 4
Consumption	-0.0048	0.0385	-0.0179	0.4449
Investment	0.0695	0.0925	-0.0297	0.3853
Inflation	0.2563	0.1517	0.4053	0.1026
Federal Funds Rate	0.4491	0.3866	0.3717	0.0304

Figure 2.1: Forecast Errors

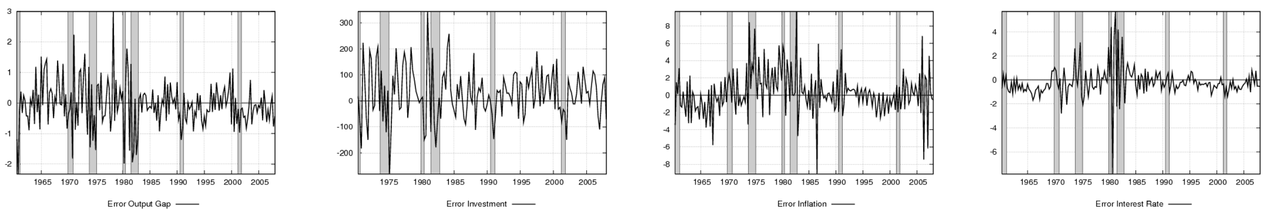
Case 1: Rational Expectations

Consumption Investment Inflation Fed Funds



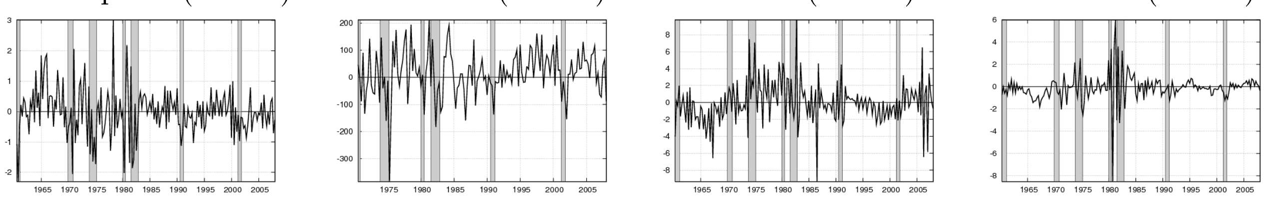
Case 2: Learing with RE Initial Conditions

Consumption (0.9467) Investment (0.8525) Inflation (0.9235) Fed Funds (0.9512)



Case 3: Learing with RE Initial Conditions, Shocks Unobservable

Consumption (0.9922) Investment (0.8598) Inflation (0.6517) Fed Funds (0.9474)



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

Consumption (0.7003) Investment (0.4776) Inflation (0.6457) Fed Funds (0.6882)

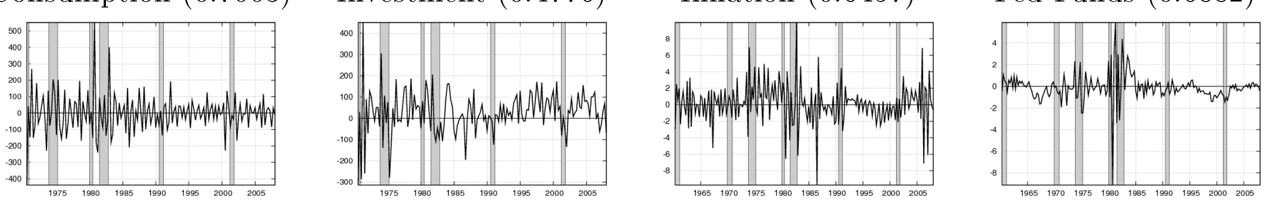


Figure 2.2: Out of Sample Multiperiod Forecast Errors

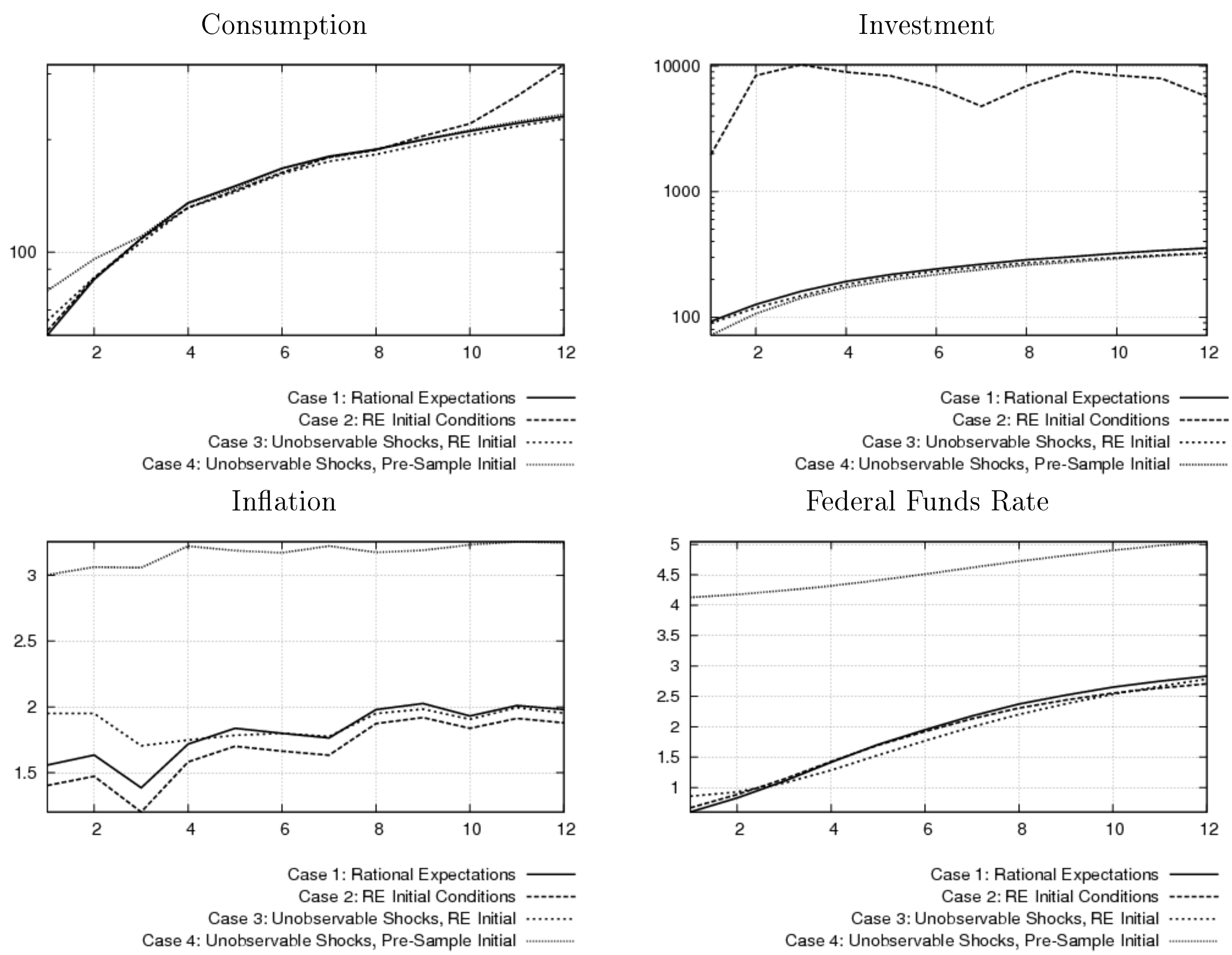
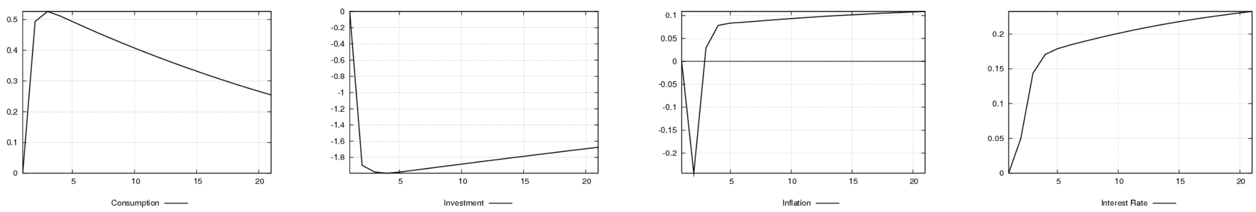


Figure 2.3: Preference Shock Impulse Responses

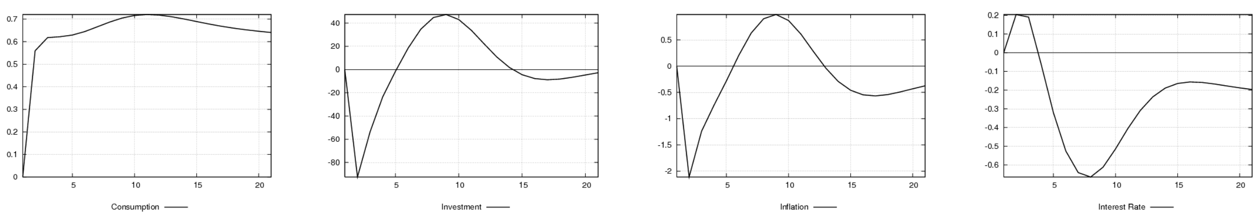
Case 1: Rational Expectations

Consumption Investment Inflation Interest Rate



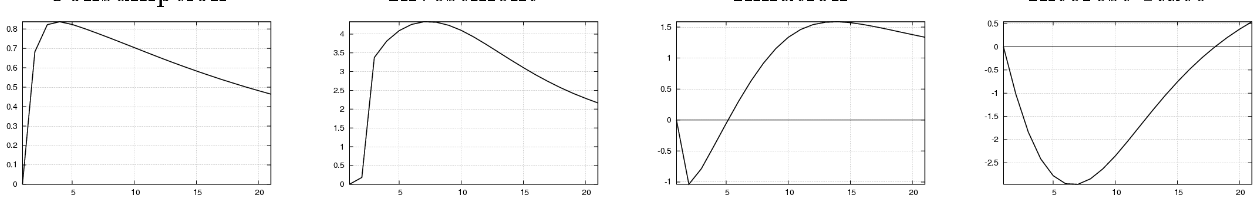
Case 2: Learning with RE Initial Conditions

Consumption Investment Inflation Interest Rate



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Consumption Investment Inflation Interest Rate



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

Consumption Investment Inflation Interest Rate

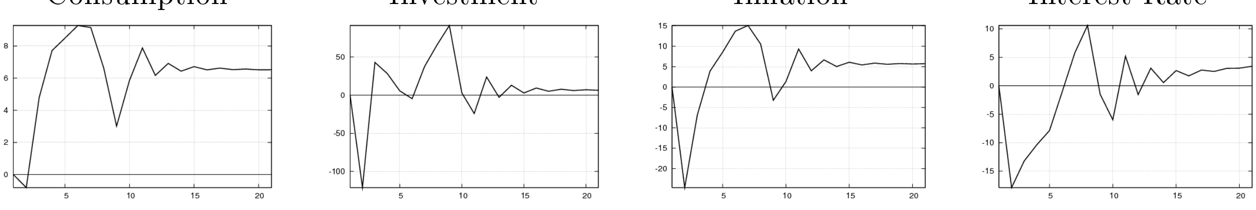
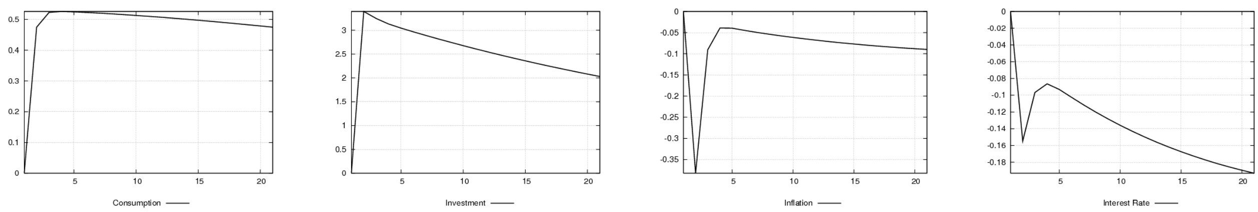


Figure 2.4: Technology Shock Impulse Responses

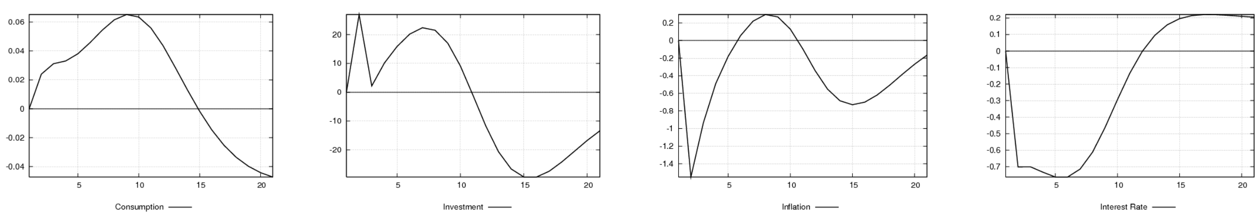
Case 1: Rational Expectations

Consumption	Investment	Inflation	Interest Rate



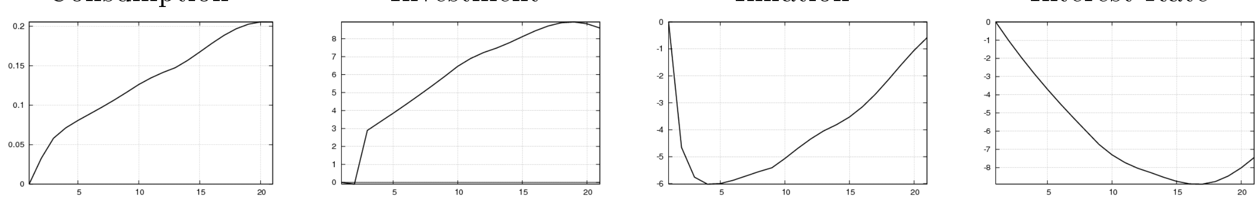
Case 2: Learning with RE Initial Conditions

Consumption	Investment	Inflation	Interest Rate



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Consumption	Investment	Inflation	Interest Rate



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

Consumption	Investment	Inflation	Interest Rate

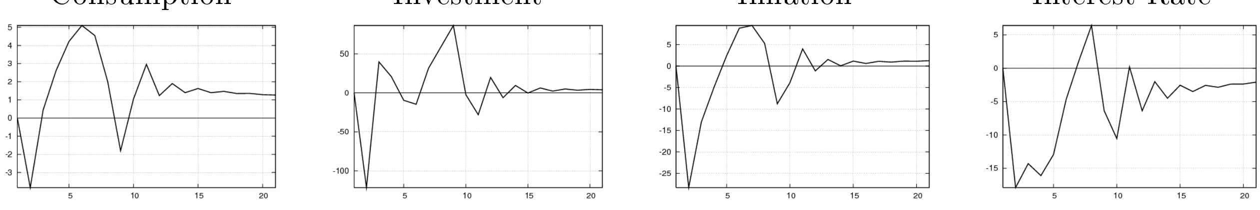


Figure 2.5: Investment Shock Impulse Responses

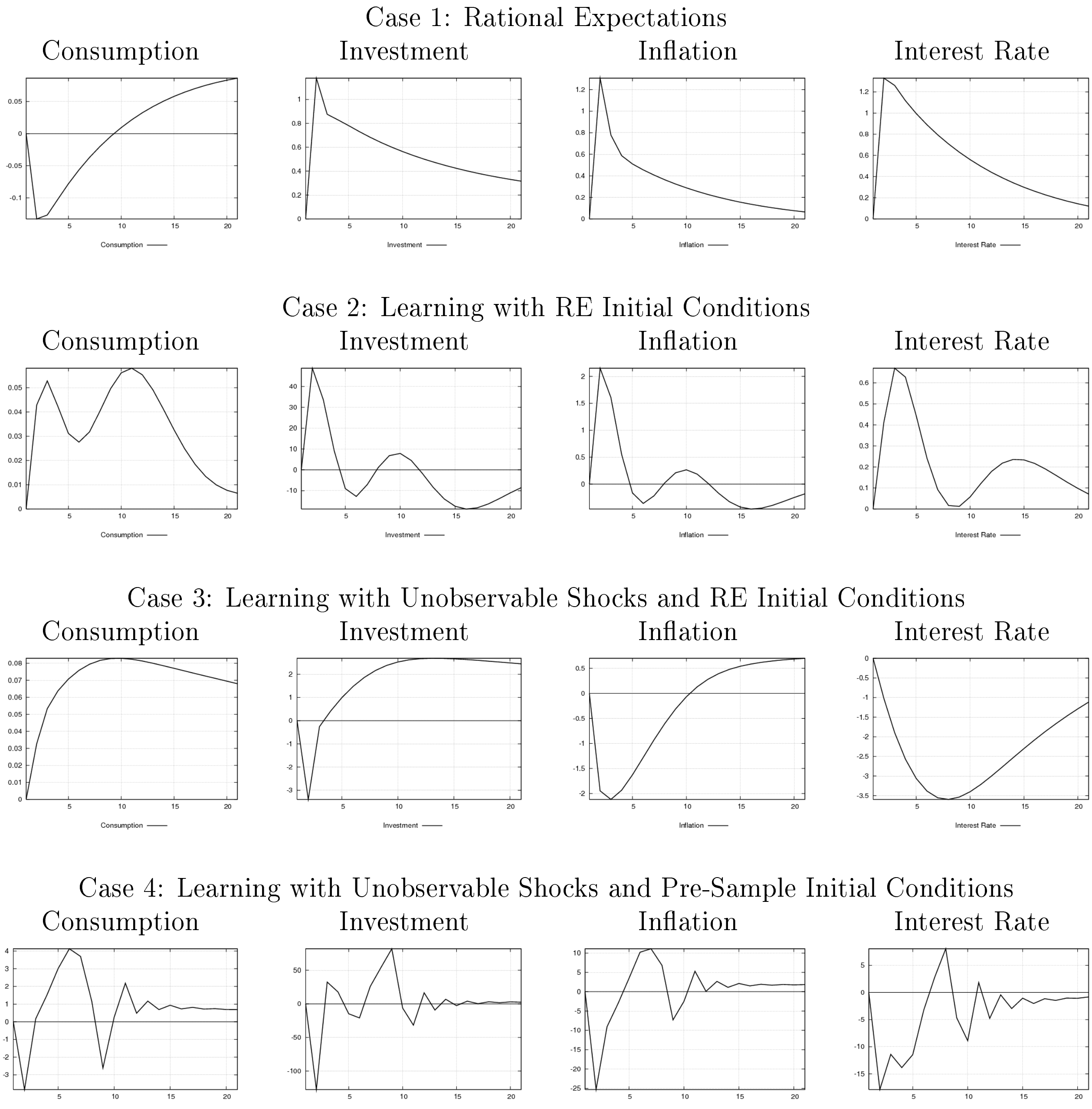
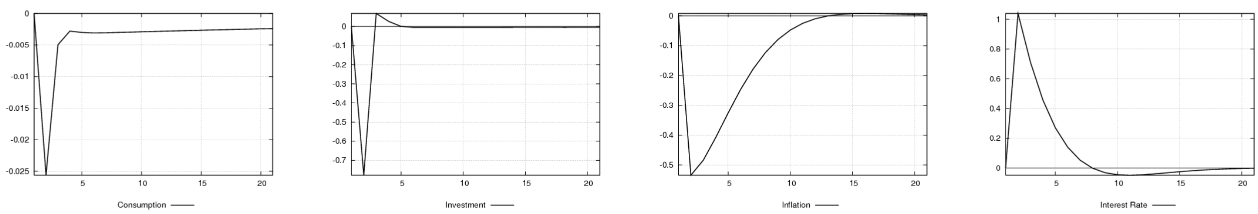


Figure 2.6: Monetary Policy Shock Impulse Responses

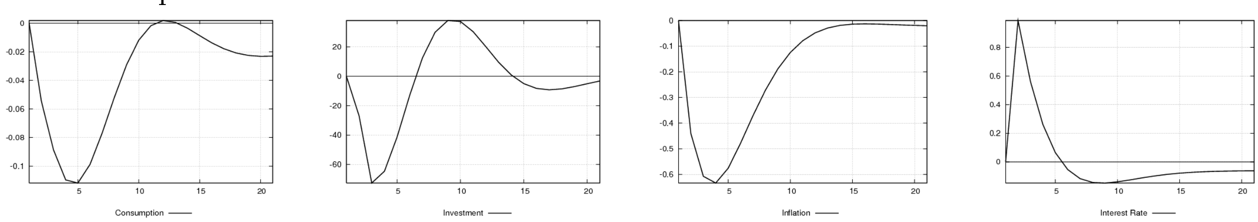
Case 1: Rational Expectations

Consumption Investment Inflation Interest Rate



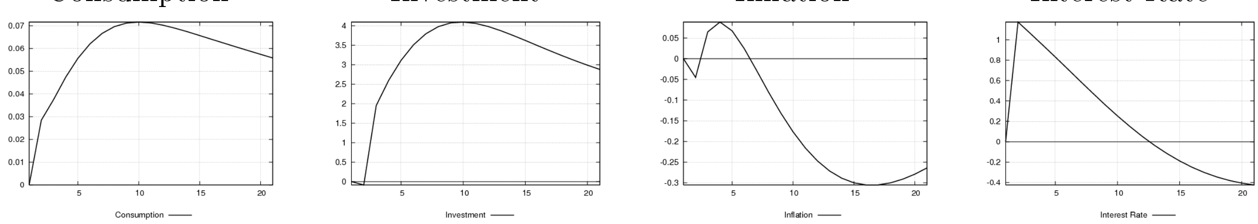
Case 2: Learning with RE Initial Conditions

Consumption Investment Inflation Interest Rate



Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Consumption Investment Inflation Interest Rate



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

Consumption Investment Inflation Interest Rate

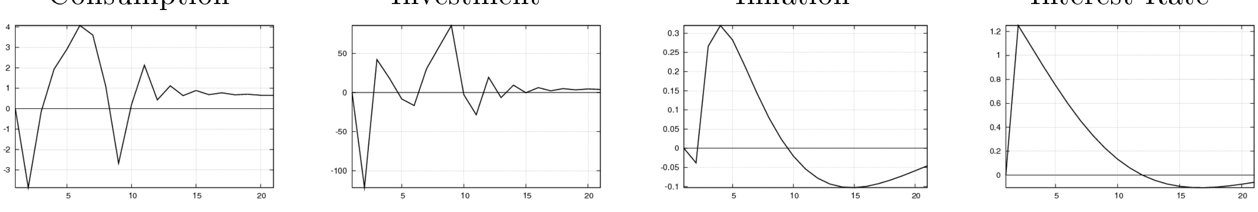
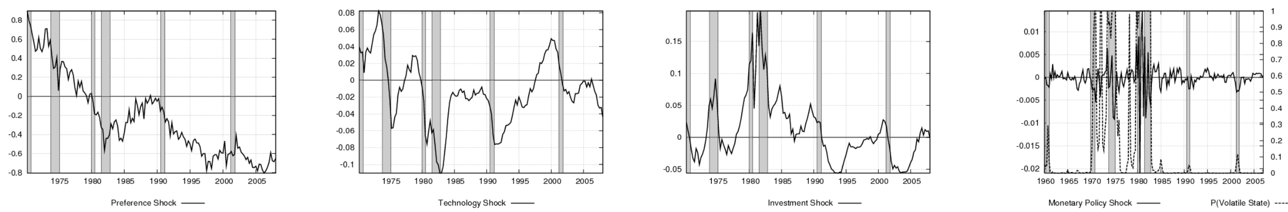


Figure 2.7: Smoothed Estimates of Structural Shocks

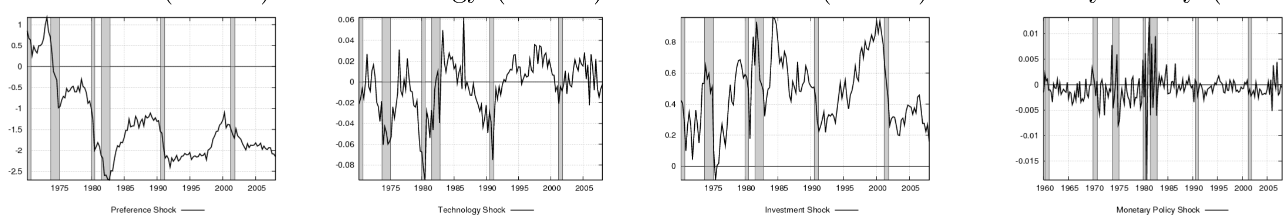
Case 1: Rational Expectations

Preference | Technology | Investment | Monetary Policy



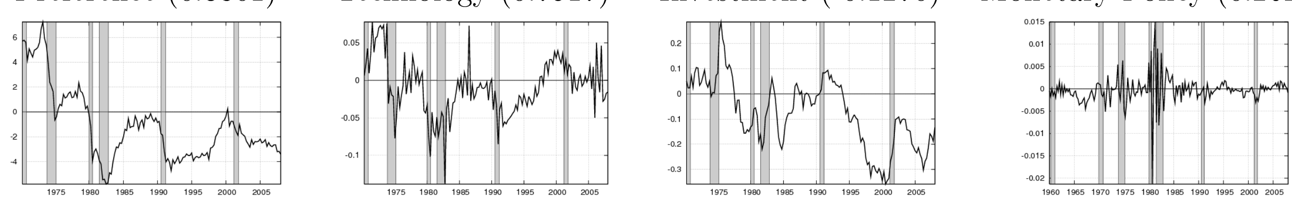
Case 2: Learning with RE Initial Conditions

Preference (0.8708) | Technology (0.0826) | Investment (0.5076) | Monetary Policy (0.7246)



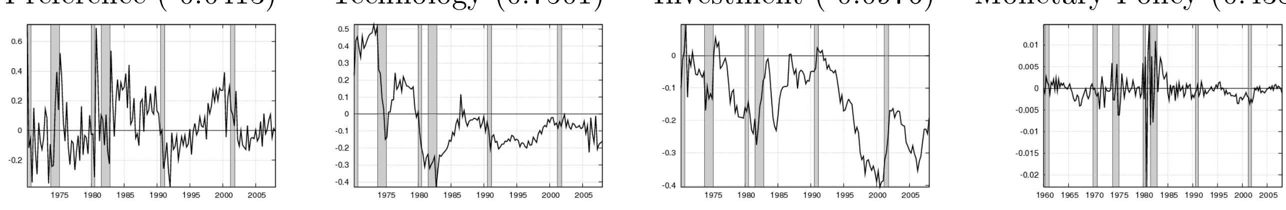
Case 3: Learning with Unobservable Shocks and RE Initial Conditions

Preference (0.8301) | Technology (0.7917) | Investment (-0.1270) | Monetary Policy (0.2623)



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

Preference (-0.0413) | Technology (0.7501) | Investment (-0.0976) | Monetary Policy (0.4380)



Chapter 3

Regime Switching, Learning, and the Great Moderation

Abstract: This paper examines the "bad luck" explanation for changing volatility in U.S. inflation and output when agents do not have rational expectations, but instead form expectations through least squares learning with an endogenously changing learning gain. It has been suggested that 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. 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. To assess whether learning can explain the Great Moderation, the New Keynesian model with volatility regime switching and dynamic gain learning is estimated by maximum likelihood. The results show that learning does lead to lower variances for the shocks in the volatile regime, but changes in regime is still significant in differences in volatility from the 1970s and after the the 1980s.

3.1 Introduction

Figure 3.1 shows plots of the U.S. output gap, the percentage difference between the actual level of real GDP and potential GDP, and the inflation rate, as measured by the annualized quarterly growth rate of the consumer price index. Aside from standard business cycle fluctuations, the data exhibits prolonged periods of differing degrees of volatility. Output and inflation are especially volatile during the 1970s and early 1980s, and there has been a subsequent decline in volatility since that period. Kim and Nelson (1999a) estimate that since the first quarter of 1984, there has been a permanently smaller difference between the growth rate of output during expansions and during recessions.

Macroeconomics has had difficulty explaining this “Great Moderation”, as it has come to be called. The leading explanations for the change in volatility fall into two groups: good vs. bad monetary policy, and good vs. bad luck. Using a standard New Keynesian model, 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, the Federal Reserve 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 led to greater volatility.

Many studies find that a change in monetary policy is not enough to explain the change in volatility. Sims and Zha (2006) find that evidence that changes in U.S. volatility is better explained by changes in luck than changes in monetary policy. Using a structural vector autoregression model with minimal identification restrictions so that it can possibly encompass many sorts of linear dynamic macroeconomic models with a monetary policy rule, they find the best fitting model is one in which there are no regime changes in the coefficients describing monetary policy or economic behavior, and there are only regime changes in the variance of the exogenous shocks. Stock and Watson (2003) similarly find that improved monetary policy accounts for only a small part of the slowdown in macroeconomic volatility since 1984.

A third explanation that is just recently receiving some attention in the literature is that expectations that evolve according to least squares learning can lead to changing periods of volatility. Under least squares learning, agents do not know the parameters that govern the economy. Not being able to form rational expectations, they estimate least squares autoregressions using past data, and use this econometric model to forecast future variables. Milani (2005) estimates a New Keynesian model with least squares learning and when splitting the sample at the same dates as Lubik and Schorfheide (2004), finds there is little evidence of a change in monetary policy and estimates the federal funds rate did adjust by more than one-to-one with inflation throughout the sample.

Primiceri (2006) demonstrates how learning on the part of the central bank can be perceived as bad monetary policy. In an empirically founded model of unemployment rate and inflation determination he shows how mis-perceptions by the central bank about the natural rate of unemployment and the degree of persistence of inflation led to a bad policy prescription during the 1970s, and therefore excessively volatile unemployment and inflation. As time progressed and more data became available, the Federal Reserve learned its mistakes and eventually provided better policy, leading to a slowdown in macroeconomic volatility after the mid-1980s. Stock and Watson (2003) provide some evidence for such an explanation when they demonstrate that univariate least squares forecasts have been more precise during less volatile periods.

Unlike previous papers in the learning literature, this paper estimates a model that combines the bad luck explanation along with learning to determine whether learning leads to a different prediction on the amount of bad luck needed to generate the changes in volatility seen in U.S. data. Previous papers in the learning literature have allowed volatility to be affected only by learning dynamics, so it is not clear whether these explanations trump the bad luck explanation. Changes in luck is modeled by assuming the variances of exogenous structural shocks are determined by a regime switching process with two states. One state is characterized by large variances of the shocks and the other has relatively smaller variances. Being in the volatile state is considered "bad luck" as the states evolve exogenously according to a Markov chain were the current state is only dependent on the previous state, and the probabilities of switching between states is exogenous.

Agents do not have rational expectations, and are therefore completely unaware of any regime switching processes. Agents do remain suspect that structural changes may occur, but they do not have any knowledge about what types of structural changes are possible, or any idea with what probabilities structural changes can occur. Expectations therefore evolve according to a process similar to the Marcet and Nicolini (2003) framework where the weight agents give to older observations is endogenous. Specifically, agents begin using a decreasing learning gain which consistent with forming ordinary least squares forecasts and having no suspicion of structural change. 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 have more likely occurred after a structural break.

While a constant learning gain is theoretically capable of producing time varying volatility, in Murray (2008) 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 using parameters estimated with U.S. data that includes the endogenously changing learning gain similar to that suggested by Marcet and Nicolini (2003) and finds that this type of learning can produce heteroskedasticity in output and inflation that rational expectations simulations cannot. Moreover, his estimates imply that during most of the 1970s decade, agents suspected structural change and therefore used a high learning gain.

This paper builds on Milani's analysis by also allowing for recurrent regime changes in the volatility of the structural shocks. Estimation of this model decomposes the changes in volatility into changes in the volatility of structural shocks and endogenous changes in the learning gain. The main finding is that learning indeed leads to much smaller variances of structural shocks in the volatile regime. However, changes in regime are still significant and the learning frameworks and rational expectations framework make similar predictions for the periods in U.S. history when the economy was in the volatile regime. Despite the ability for learning to explain much of the volatility, the change in the dynamic gain appears to play an insignificant role, and the rational expectations model dominates the learning models in terms of its fit to the data.

The paper proceeds as follows. Section 2 presents the New Keynesian framework and regime switching process. Section 3 describes the learning process and how constant gain learning and dynamic gain learning can generate time varying volatility. Section 4 describes the data and estimation procedure. Section 5 presents the estimation results and interprets the findings, and Section 6 concludes.

3.2 Model

The crucial extensions of learning with an endogenously changing gain and a Markov chain process for changes in volatility are incorporated into a standard New Keynesian model of output and inflation determination and monetary policy. This section describes the rational expectations version New Keynesian model and the next section introduces learning in the linearized version of the model.[1]

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 has 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

[1] This methodology is perhaps the most common means for putting learning into macroeconomic models, but Preston (2005) demonstrates that since the least squares expectations operator does not follow the law of iterated expectations, this method is not consistent with learning inherent in the microfoundations of the model.

intermediate goods. The intermediate goods are imperfect substitutes for each other in production, therefore intermediate 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.

3.2.1 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. However, given a perfect asset market, 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,

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

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 extensions for both the United States and Euro area and find point estimates for the degree of habit formation very close to unity. 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) shows that under learning the estimate for the degree of habit formation falls close to zero.

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

$$\hat{\lambda}_t = E_t\hat{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1}, \tag{3.1}$$

where a hat indicates the percentage deviation of the variable from its steady state.[2] Here, $\hat{\lambda}_t$ is the marginal utility of real income, given by,

$$\hat{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\hat{c}_{t+1} - \sigma(1+\beta\eta^2)\hat{c}_t + \sigma\eta\hat{c}_{t-1}\right] + \left(\hat{\xi}_t - \beta\eta E_t\hat{\xi}_{t+1}\right). \tag{3.2}$$

3.2.2 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,

$$y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di\right]^{\frac{\theta}{\theta-1}}, \tag{3.3}$$

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,

$$y_t(i) = \left[\frac{p_t(i)}{p_t}\right]^{-\theta} y_t, \tag{3.4}$$

[2] 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.

where $p_t(i)$ is the price of intermediate good i and p_t is the price of the final good. Substituting equation (3.4) into (3.3) leads to a consumption price index that holds in equilibrium,

$$p_t = \left[\int_0^1 p_t(i)^{1-\theta} di \right]^{\frac{1}{1-\theta}}. \tag{3.5}$$

Each intermediate good is produced according the constant returns to scale production function $y_t(i) = \zeta_t n_t(i)$, where ζ_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,

$$\hat{\psi}_t(i) = \frac{1}{\mu}\hat{y}_t(i) - \hat{\lambda}_t - \left(\frac{1}{\mu}+1\right)\hat{\zeta}_t. \tag{3.6}$$

Summing equation (3.6) across all firms leads to the average marginal cost in the economy,

$$\hat{\psi}_t = \frac{1}{\mu}\hat{y}_t - \hat{\lambda}_t - \left(\frac{1}{\mu}+1\right)\hat{\zeta}_t. \tag{3.7}$$

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:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_t(i)\pi^*_{t+T}}{p_{t+T}} \right) y_{t+T}(i) - \Psi\left[y_{t+T}(i)\right] \right\}, \tag{3.8}$$

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 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 (3.5), leads to the log-linear Phillips equation,

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{\mu(1-\omega)(1-\omega\beta)}{\omega(\mu+\theta)}\hat{\psi}_t\right] \tag{3.9}$$

3.2.3 Fully Flexible Prices

To take this model to data on the output gap it is convenient to rewrite the model in terms of the difference from the outcome in which there are no nominal rigidities. Let $\tilde{y}_t = \hat{y}_t - \hat{y}_t^f$ and $\tilde{\lambda}_t = \hat{\lambda}_t - \hat{\lambda}_t^f$ denote the percentage deviation of output and marginal utility from their fully flexible price outcome. Under fully flexible prices the linearized Euler equation, (3.1), and marginal utility of income, (3.2), still hold. Using these conditions and imposing goods market equilibrium condition implies,

$$\tilde{\lambda}_t = E_t\tilde{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1} - r_t^n, \tag{3.10}$$

$$\tilde{\lambda}_t = \frac{1}{(1-\beta\eta)(1-\eta)}\left[\beta\eta\sigma E_t\tilde{y}_{t+1} - \sigma(1+\beta\eta^2)\tilde{y}_t + \sigma\eta\tilde{y}_{t-1}\right], \tag{3.11}$$

where r_t^n is the percentage deviation of 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,

$$r_t^n = \rho_n r_{t-1}^n + \epsilon_{n,t}, \tag{3.12}$$

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

When prices are fully flexible, $\omega = 0$ in the maximization problem given in equation (3.8). It can be shown in this case that the first order condition implies the marginal cost is identical for every firm and always constant. Under fully flexible prices equation (3.7) implies,

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

One can solve this equation for $\hat{\zeta}_t$ and substitute that back into the equation for marginal cost (3.7). Plugging this expression for marginal cost into equation (3.9) yields the following Phillips curve in terms of the output gap,

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\omega(\mu+\theta)}(\tilde{y}_t - \mu\tilde{\lambda}_t)\right]. \tag{3.13}$$

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,

$$\pi_t = \frac{1}{1+\beta\gamma}\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \kappa(\tilde{y}_t - \mu\tilde{\lambda}_t) + u_t\right], \tag{3.14}$$

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

$$u_t = \rho_u u_{t-1} + \epsilon_{u,t}, \tag{3.15}$$

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

3.2.4 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 of the form,

$$\hat{r}_t = \rho_r\hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi E_t\pi_{t+1} + \psi_y E_t\tilde{y}_{t+1}\right) + \epsilon_{r,t} \tag{3.16}$$

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 expected inflation, $\psi_y \in (0,\infty)$ is the feedback on the interest rate to the expected output gap, and $\epsilon_{r,t}$ is an independently and identically distributed exogenous monetary policy shock.

3.2.5 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 natural interest rate shock $\epsilon_{n,t}$, the cost push shock, $\epsilon_{u,t}$, and the monetary policy shock $\epsilon_{r,t}$ are determined by two states. Let state L, denoted by s_L, be the state where the shocks have low volatility, and state H, denoted by s_H, be the state where the shocks have high volatility.

The variances of the structural innovations in a given state are independently normally distributed with mean zero and variances given by,

$$Var\left[\epsilon_t(s_t)\right] = \left\{ \begin{array}{l} \begin{bmatrix} \sigma^2_{n,L} & 0 & 0 \\ 0 & \sigma^2_{u,L} & 0 \\ 0 & 0 & \sigma^2_{r,L} \end{bmatrix}, \text{ if } s_t = L \\ \begin{bmatrix} \sigma^2_{n,H} & 0 & 0 \\ 0 & \sigma^2_{u,H} & 0 \\ 0 & 0 & \sigma^2_{r,H} \end{bmatrix}, \text{ if } s_t = H \end{array} \right\}, \tag{3.17}$$

where $\epsilon_t(s_t)' = [\epsilon_{n,t}(s_t)\ \epsilon_{u,t}(s_t)\ \epsilon_{r,t}(s_t)]$, and the variances in the high volatility state are greater than or equal to the variances in the low volatility state.

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 $j \in \{L, H\}$. This implies that the probability of moving from state i in $t-1$ to state j in time t, where $i \neq j$ is given by $1 - p_j$. The state then evolves according to the transition matrix,

$$P = \begin{bmatrix} p_L & 1-p_L \\ 1-p_H & p_H \end{bmatrix}. \tag{3.18}$$

Let $S_t' = [P(s_t = L)\ P(s_t = H)]$ denote the probability of being in each state at any given time t. The transition matrix assumes the state evolves according to,

$$E_t S_{t+1} = P S_t \tag{3.19}$$

Notice this regime switching framework makes the somewhat restrictive assumption that all structural shocks are always in the same regime. A less restrictive assumption would be that one or more shocks could be in one regime while one or more others could be in another regime. Such a setup would require specifying a transition matrix for each structural shock. If one assumed that the different shocks transitioned between regimes independently of one another, this would introduce four more parameters to be estimated. If one wanted to generalize the process so that there is some dependence, this would involve even more parameters. To keep the number of parameters to estimate tractable, and to avoid over-fitting the data, I assume that all shocks are in the same regime in a given time period.

3.3 Learning

Instead of having rational expectations, agents form expectations by estimating least squares regression models, where expectations of future output and inflation are given by the forecasts from these models. Agents are assumed to have no knowledge of the structural form of the economy, the parameters that govern the economy, or the regime switching process. They do know the reduced form of the economy follows a VAR(1) process, and they use this model and past data to form their forecasts.

The model in the previous section only allows for structural changes in the volatility in the shocks, and since the model is linearized, agents are indifferent to the additional risk. Even so, I suppose that agents suspect that structural changes of unknown types are possible so they may decide to give more weight to recent observations in their estimation procedure. One way to model this is to use constant gain learning, which is consistent with agents using forecasts based on weighted least squares estimates, 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.

Authors such as Sargent (1999) and Evans and Honkapohja (2001) have suggested that constant gain learning is a natural way to model expectations in a real world where structural changes, perhaps large or small, are always possible. The drawback of constant gain learning is that it implies agents always have the same level of suspicion for structural changes, regardless of recent macroeconomic activity or the size of recent forecast errors. 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, causing learning dynamics to disappear completely in the long run. Learning with OLS also implies that agents believe structural changes that should impact their decisions are impossible.

Constant gain learning can in theory lead to time varying volatility of expectations, and therefore time varying volatility for output and inflation. This affect depends on the size of the constant learning gain, but for empirically plausible values, this type of learning fails to deliver very big effects. Williams (2005) finds this to be the case in a simple model with simulated data.

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 recent structural changes. As time progresses they form their expectations using a decreasing learning gain consistent with OLS. Agents will suspect a structural change may have occurred if recent forecast errors are larger than historical averages. When this happens, agents switch to a larger learning gain, which puts more weight on the most recent observations, ones that are believed to have more likely occurred after a structural change. As long as forecast errors are large, the learning gain remains at this high point. When forecast errors start becoming small, the learning gain decreases at a rate consistent with OLS. I explain in more detail the endogenous learning gain process 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.

3.3.1 Least Squares Learning

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

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi z_t, \tag{3.20}$$

$$z_t = A z_{t-1} + \epsilon_t(s_t) \tag{3.21}$$

where E_t^* denotes possibly non-rational expectations, x_t is a vector of minimum state variables, and z_t is a vector of structural shocks. For the New Keynesian model, $x_t = [\tilde{y}_t \ \pi_t \ \hat{r}_t \ \tilde{\lambda}_t]'$ and $z_t = [r_t^n \ u_t \ \epsilon_{r,t}]'$. The minimum state variable solution of the model implies the rational expectation for x_{t+1} is given by,

$$E_t x_{t+1} = G x_t + H E_t z_{t+1}, \tag{3.22}$$

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. Agents that learn do not know the the parameters that govern the economy, but do use the reduced form of the economy for their forecasting model. Agents' information sets are restricted only to past data on x_t, so they are unable to collect data on past structural shocks to estimate matrix H. Therefore, agents collect past data on x_t to form least squares estimates for the non-zero columns of G.

Agents do know what columns of G are equal to zero, and therefore do not use the associated variables as explanatory variables in their regression. In terms of the New Keynesian model in the previous section, the only non-zero column of G is that which multiplies past marginal utility of income, $\tilde{\lambda}_t$. When there is a positive degree of habit formation, not only are expectations of next period's output important for consumers' decisions, so is next period's future marginal utility of income which involves a two period ahead forecast for output. Since the only sources of persistence in the model are on output (habit formation), inflation (price indexation), and the interest rate (monetary policy smoothing), these are the only variables whose lags agents use as explanatory variables.

The timing in which agents make expectations and decisions in a given period is as follows. At the beginning of period t agents wake up with data realized through period $t-1$. They collect this data and use the least squares estimate for G 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 become available as data to the agents, and the process begins again.

There is no constant term in the general form of the model, equation (3.20), or in the rational expectation, given in equation (3.22), since all variables are expressed in terms of percentage deviations from the steady state or flexible price outcome. Since agents are not endowed with information about the parameters of the model, it is realistic to suppose that agents also estimate a constant term in equation (3.22). Let $\hat{G}_t^*$ denote agents' time t estimate for the non-zero columns of matrix G and a column for a constant term so that $\hat{G}_t^* = [\hat{g}_t \ \hat{G}_t^{NZ}]$, where $\hat{g}_t$ is the time t estimate of the constant term and $\hat{G}_t^{NZ}$ is the time t estimate for the non-zero columns of G.

If agents use OLS, then,

$$\left(\hat{G}_t^*\right)' = \left(\frac{1}{t-1}\sum_{\tau=1}^{t-1} x_{\tau-1}^* {x_{\tau-1}^*}'\right)^{-1} \left(\frac{1}{t-1}\sum_{\tau=1}^{t-1} x_{\tau-1}^* x_\tau'\right), \tag{3.23}$$

where ${x_t^*}' = [1 \ x_t^{NZ'}]$ is the vector of explanatory variables agents use. Agents form the expectation,

$$E_t^* x_{t+1} = \hat{g}_t + \hat{G}_t E_t^* x_t = (I + \hat{G}_t)\hat{g}_{0,t} + \hat{G}_t^2 x_{t-1}, \tag{3.24}$$

where $\hat{G}_t$ denotes the time t estimate for G obtained from $\hat{G}_t^*$ with the zero columns filled back in. The least squares estimate for $\hat{G}_t^*$ can be rewritten in the following recursive form:

$$\hat{G}_t^* = \hat{G}_{t-1}^* + g_t(x_{t-1} - \hat{G}_{t-1}^* x_{t-2}^*) x_{t-2}^{*}{}' R_t^{-1}, \tag{3.25}$$

$$R_t = R_{t-1} + g_t(x_{t-2}^* x_{t-2}^{*}{}' - R_{t-1}), \tag{3.26}$$

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 (3.25) is the realized forecast error for the previous estimate $\hat{G}_{t-1}^*$. The degree to which agents adjust their expectations depends on the size of this forecast error, the variance of the estimated coefficients, captured by the inverse of matrix R_t, and 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 the degree to which new observations can affect expectations is always the same.

Substituting the expectation in equation (3.24) into the structural form, (3.20), leads to the following evolution for x_t,

$$x_t = \Omega_0^{-1}\Omega_2(I + \hat{G}_t)\hat{g}_{0,t} + \Omega_0^{-1}\left(\Omega_1 + \Omega_2 \hat{G}_t^2\right) x_{t-1} + \Omega_0^{-1}\Psi z_t, \tag{3.27}$$

where $\hat{g}_{0,t}$ and $\hat{G}_t$ are determined by the learning process in equations (3.25) and (3.26).

This form illustrates how learning with a positive learning gain can lead to time-varying volatility for x_t even if with a constant variance for z_t. Unlike standard rational expectations models, the constant term and matrix multiplying x_{t-1} are time-varying. The magnitude of these matrices depends on the size of the learning gain and the size of agents' forecast errors. Time variation in these matrices causes time variation in the volatility of x_t. Note this is true for any non-zero value for the learning gain g_t. Even a constant learning gain can technically generate time-varying

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

volatility, but time variation in the learning gain has been suggested by Marcet and Nicolini (2003) and Milani (2007) to better explain such phenomenon.

3.3.2 Dynamic Gain Learning

Under a mixed learning framework, 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 evolve according to,

$$\alpha_t = \left\{ \begin{array}{ll} \alpha_{t-1} + 1 & \text{if } \frac{1}{J}\sum_{j=1}^{J}\frac{1}{n}\sum_{v=1}^{n}\left|x_{t-j}(v) - \hat{G}^*_{t-j}(v)x^*_{t-j-1}\right| < \nu_t \\ \alpha & \text{otherwise} \end{array} \right\} \tag{3.28}$$

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 $\hat{G}^*_{t-j}$, 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 may change their learning gain looking at only the most recent forecast error. I fix $J = 8$, so for quarterly data agents examine the forecast errors from the most recent two years and adjust the learning gain if the average forecast error during this time is too large. The variable $\nu \in (0, \infty)$ is the threshold for how large the forecast errors must be to induce agents to increase their learning gain. Similar to Milani (2007), this threshold is set equal to the average size of forecast errors up through date $t-1$, which is given by,

$$\nu_t = \frac{1}{t-1}\sum_{j=1}^{t-1}\frac{1}{n}\sum_{v=1}^{n}\left|x_{t-j}(v) - \hat{G}^*_{t-j}(v)x^*_{t-j-1}\right|.$$

Since forecast errors for each variable v is given as a percentage deviation from the steady state or potential, they are added up over all the variables that agents forecast. This learning mechanism can nest the special cases when agents always use OLS or always use a constant gain. To restrict agents to always use OLS, ν_t can be set to zero for all t. To restrict agents to always use a constant gain, ν_t can be set to infinity for all t.

This learning mechanism introduces one additional parameter to estimate jointly with the parameters of the New Keynesian model and regime switching process, the threshold learning gain, $g \equiv 1/\alpha$.

3.4 Estimation

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 conforms to a state-space representation with Markov-switching in the variance of the error term and is estimated using the Kim and Nelson (1999b) technique for combining the Kalman filter that evaluates a state-space model with the Hamilton (1989) filter for evaluating Markov-switching processes.

3.4.1 Maximum Likelihood Procedure

The state side of the model is given by equations (3.27) and (3.21). Let GAP_t denote data on the output gap, INF_t denote data on inflation, and FF_t denote data on the Federal Funds rate. The observation equations are given by,

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

The state variables are multiplied by 100 to convert the decimals into percentages, and the inflation rate and federal funds rate are 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 $r^* = 400(1 - 1/\beta)$.

The log-likelihood is maximized with respect to the threshold learning gain, g, the Markov-switching probabilities, p_h and p_L, the parameters of the New Keynesian model, and the variances of the structural shocks for each regime. The discount factor is not estimated but instead fixed to $\beta = 0.99$ which implies a steady state real interest rate of about 4%. The elasticity of substitution between intermediate goods, θ, and the degree of price flexibility, ω, appear multiplicatively in the Phillips curve (3.13) and so only the reduced form coefficient, κ, is estimated. Before revealing the estimation results, it is first necessary to specify how initial conditions for the learning process are set.

3.4.2 Initial Conditions

Aside from standard initial conditions for the Kalman filter and Hamilton filter, it is necessary to specify initial conditions for the learning process given in equations (3.25) and (3.26). How values are set for initial expectation matrices, $\hat{g}_0$, $\hat{G}_0^*$, and R_0, can have a dramatic effects on the estimation results. Despite this dependence, there is little general consensus for how initial expectations should be specified.

Williams (2005) 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.

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 (2005, 2007) sets the initial conditions for the learning matrices equal to VAR(1) estimates from de-meaned pre-sample data. Similarly, I estimate the appropriate regressions with pre-sample data from 1954:Q2 through 1959:Q4. In the

New Keynesian framework, agents estimate four regression models for the following dependent variables: output gap, inflation rate, interest rate, and marginal utility of income gap. Each of these variables depends on lagged output gap, inflation rate, and interest rate. The data must first be transformed into the same terms as the state vector, x_t. For the output gap, inflation rate, and interest rate this is done according to:

$$\tilde{y}_t = \frac{1}{100}GAP_t$$
$$\pi_t = \frac{1}{400}(INF_t - \pi^*),$$
$$\hat{r}_t = \frac{1}{400}(FF_t - r^* - \pi^*).$$

Expectations for these first three variables in the state vector is found by estimating a VAR(1) on $[\tilde{y}_t\ \pi_t\ \hat{r}_t]'$. Data for the marginal utility of income gap is found by plugging into equation (3.11) data for the output gap, lagged output gap, and expected future output gap predicted by the VAR(1). Expectations for $\tilde{\lambda}_t$ are then found by regressing this simulated data on lagged $[\tilde{y}_t\ \pi_t\ \hat{r}_t]'$.

3.5 Results

To analyze how learning and regime switching in volatility are related, the New Keynesian model is estimated under rational expectations, constant gain learning, and dynamic gain learning gain. Table 3.6 presents the parameter estimates for each case. Under constant gain learning, the estimate for learning gain is essentially equal to zero. This implies that expectations do not evolve through the sample. Even so, the predictions under constant gain learning are not the as same rational expectations, since the coefficient matrices agents use to form expectations are different. The expectation matrices for the learning case are initialized to pre-sample VAR(1) results, which do not coincide with rational expectations. The estimated threshold gain under dynamic gain learning has a very small point estimate, $g = 0.0045$, but it is statistically significantly different from zero. This implies that expectations do evolve over time, but the rate at which agents learn is very small.

Comparing the standard deviation of the shock processes illustrates how expectations explain macroeconomic volatility. There is little difference in the predictions for the volatility of the cost push and monetary policy shocks, but there is a substantial difference in the volatility for the natural rate shock. For both the low and high volatility regime, the estimate for the variance of the natural rate shock under rational expectations is almost twice as high as under constant gain learning, and almost four times higher than under dynamic gain learning. This implies that using the VAR(1) on pre-sample data to specify expectations explains much of the volatility in output, but evolving expectations with dynamic learning gain explains even more. The persistence of the natural rate shock is also somewhat larger under rational expectations ($\rho_n = 0.8705$), than under dynamic gain learning ($\rho_n = 0.7484$) or constant gain learning ($\rho_n = 0.6936$).

Other parameter estimates that differ across models include the intertemporal elasticity of substitution and the Phillips curve slope. The elasticity of substitution is approximately $\sigma = 0.0073$ under rational expectations, $\sigma = 0.2560$ under dynamic gain learning, and $\sigma = 0.1824$ under constant gain learning. This implies the intertemporal consumption trade-off is much more sensitive to the expected real interest rate with the initial expectations for the learning processes than under rational expectations. Moreover, dynamic gain learning leads to an even higher estimate than the zero constant gain, although the difference is not statistically significant.

The estimates for the Phillips curve coefficient reveals how learning dynamics can alter the prediction for the degree of price flexibility. The coefficients under rational expectations and constant gain learning are both very close to zero, which implies a very small degree of price flexibility. However, under dynamic gain learning, the only framework in which expectations are evolving, the Phillips curve coefficient is much larger, implying a greater degree of price flexibility.

The Markov probabilities are very similar across the models. Both the low and high volatility regimes are very persistent which implies changes in luck is still very important in explaining changes in volatility in the sample, regardless of how expectations are formed. Figure 3.2 shows plots the smoothed estimate for the probability of being in the high volatility regime throughout the sample for each of the three models. The middle panel also shows the evolution of the dynamic learning gain during the same time. All three models predict strong probabilities for being in each regime for most of the sample. During the early 1970s, middle 1970s and late 1970s and early 1980s, all models predict the economy was in the volatile regime, with brief returns to the low volatility regime between these times. Since 1985, all models predict the economy has remained in the low volatility regime. These results are consistent with previous studies such as Kim and Nelson (1999a) and Justiniano and Pimiceri (2006) which conclude changes in the volatility of exogenous shocks are significant in explaining time-varying macroeconomic volatility.

The expected number of quarters the U.S. economy has spent in the volatile regime can be found by summing the probabilities for each period over the entire sample period. Doing so reveals the economy is in the volatile regime for an expected 7.77 years under rational expectations, 9.17 years under dynamic gain learning, and 12.26 years under constant gain learning. The greater number of volatile periods predicted under the zero constant gain implies the initial conditions for expectations leads to more volatile periods, and the smaller estimate predicted under dynamic learning implies that evolving expectations may reduce the need for the number of volatile periods. Since the dynamic learning gain is so small, expectations are very slow to evolve, so this effect still does not outweigh the effect of the initial expectations.

The plot of the dynamic learning gain in the middle panel of figure 3.2 demonstrates that while the learning gain was always, there was actually little movement in the gain throughout the sample. Since the estimate for threshold learning gain is so small, there is little it can move as time progresses. The plot indicates that throughout the 1970s agents forecast errors were larger than the historical average and began to decline since 1984. The learning gain again jumped at the end of the 2001 recession and remained at the threshold level until 2003.

Figures 3.3, 3.4, and 3.5 show the smoothed estimated paths of the natural rate, cost push, and monetary policy shocks, with the probability of being in the volatile regime superimposed. The volatility of all three shocks are significantly greater in the volatile regime under each specification of the model. Comparing the natural rate shock paths shows again that rational expectations predicts much larger shocks and greater persistence than the learning models. All models indicate recessions are characterized by negative natural interest rate shocks, especially the recessions in 1974 and 1981.

The cost push shocks are very similar across models, with the largest shocks occurring during the 1970s and early 1980s. The learning models predict somewhat larger negative cost push shocks during the 1974 and 1981 recessions. The monetary policy shocks are very small throughout most of the sample with the exception of very volatile shocks during the recessions of 1974, 1979, and 1981.

Figure 3.6 shows the evolution of agents expectations for inflation and output under learning and rational expectations. In each plot, the solid line represents the smoothed estimate for the output gap and inflation, the dashed line represents expectations under learning, and the dotted line represents what the rational expectation would be with the New Keynesian parameters estimated for each learning specification. The results show that under both learning models, the expectation for next period's inflation lags slightly behind the rational expectation. This is to be expected, since learning expectations do not have access to information about the shock process z_t, but rational expectations does. The expectations for inflation under dynamic gain learning are especially close to rational expectations, while there is somewhat less volatility for inflation under the zero constant gain.

The paths of expected output under learning and rational expectations are very different. Under both learning frameworks the implied rational expectations are much more volatile. Moreover, rational expectations over-estimate the output gap throughout most of the sample, while the learning models lead to under-estimates of the output gap for much of the 1970s and 1980s and small over-estimates during the expansionary periods of the 1960s and 1990s.

Table 3.6 presents a number of criteria for comparing the relative fit of the three models. The root mean squared error (RMSE) for the output gap, inflation, and Federal Funds rate are all smallest under rational expectations, but only by a very small amount. To determine if the three models adequately explain time-varying volatility, a first order autoregression is estimated on the squared residuals for each model. Despite the regime switching process for stochastic volatility, the results indicate the variance of the residuals is still significantly positively autocorrelated for most of the cases, with the exception of the output gap under rational expectations, and the output gap and inflation under constant gain learning.

Plots of the residuals in Figures 3.7, 3.8, and 3.9 show that the largest errors for the output gap and inflation are made during the 1970s and early 1980s, and very large errors are made for the federal funds rate as Paul Volcker begins his tenure as chairman of the Federal Reserve. These failures of the model are typical for standard New Keynesian models with rational expectations and no regime changes, indicating these extensions still do not fully explain time-varying macroeconomic volatility.

3.6 Conclusion

Estimates of the New Keynesian model with dynamic gain learning and Markov-switching volatility indicate that dynamic gain learning and expectations specified by VAR(1) estimates on pre-sample data lead to much lower predictions for the variance of the natural rate shock in the low and high volatility regimes, however changes in volatility in U.S. history still depends on exogenous changes in the volatility of structural shocks. Said another way, changes in luck is still an empirically important explanation for time-varying volatility, but the degree of bad luck needed is smaller under learning. Most of this decrease in bad luck, that is the decrease in the variance of the structural shocks, is found to come from the specification of initial expectations, but some is explained by the time-variation in expectations predicted by the dynamic learning gain process. Analysis of the smoothed estimate for the evolution of the probability of being in the high volatility regime indicate the United States was in the volatile regime during much 1970s and early 1980s, a finding which is robust for rational expectations, constant gain learning, and dynamic gain learning. Nonetheless, under dynamic gain learning, expectations do evolve slowly over the sample and agents have the largest learning gain during these same periods of U.S. history.

References

Bullard, J., and J. Duffy (2004): "Learning and Structural Change in Macroeconomic Data," Federal Reserve Bank of St. Louis Working Paper 2004-016A.

Bullard, J., and S. Eusepi (2005): "Did the Great Inflation Occur Despite Poicymaker Commitment to a Taylor Rule?," Review of Economic Dynamics, 8, 324–359.

Bullard, J., G. W. Evans, and S. Honkapohja (2005): "Near Rational Exuberance," Federal Reserve Bank of St. Louis Working Paper 2004-025B.

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

Carceles-Poveda, A., and C. Giannitsarou (2005): "Adaptive Learning in Practice," Working Paper.

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.

——— (2008): "Expectations, Learning and Monetary Policy: An Overview of Recent Research," Centre for Dynamic Macroeconomic Analysis Working Paper CDMA08/02.

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. (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.

Ireland, P. N. (2004a): "A method for taking models to the data," Journal of Economic Dynamics and Control, 28, 1205–1226.

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

——— (2005): "Irrational Expectations and Econometric Practice," Federal Reserve Bank of Atlanta Working Paper 2003-22.

——— (2006): "Changes in the Federal Reserve's Inflation Target: Causes and Consequences," Working Paper.

Justiniano, A., and G. E. Pimiceri (2006): "The Time Varying Volatility of Macroeconomic Fluctuations," NBER Working Paper 12022.

Kim, C.-J. (1994): "Dynamic Linear Models with Markov-Switching," Journal of Econometrics, 60, 1–22.

Kim, C.-J., and C. R. Nelson (1999a): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," Review of Economics and Statistics, 81, 608–616.

——— (1999b): State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications. MIT 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.

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.

Mark, N. C. (2005): "Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dyanamics," NBER Working Paper 11061.

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

Milani, F. (2005): "Learning, Monetary Policy Rules, and Macroeconomic Stability," Working Paper.

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

Murray, J. M. (2008): "Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital," Indiana University mimeo.

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.

Roberts, J. M. (1995): "New Keynesian Economics and the Phillips Curve," Journal of Money, Credit and Banking, 27, 975–984.

Rotemberg, J. (1982): "Sticky Prices in the United States," Journal of Political Economy, 90, 1187–1211.

Rotemberg, J., and M. Woodford (1997): "An Optimization Based Econometric Framework for the Evaluation of Monetary Policy," In: Bernanke, B.S. and J. Rotemberg (Eds.), NBER Macroeconomics Annual. MIT Press.

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

Sims, C. (2000): "Solving Linear Rational Expectations Models," Unpublished manuscript.

Sims, C. A., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy," American Economic Review, 96, 54–81.

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.

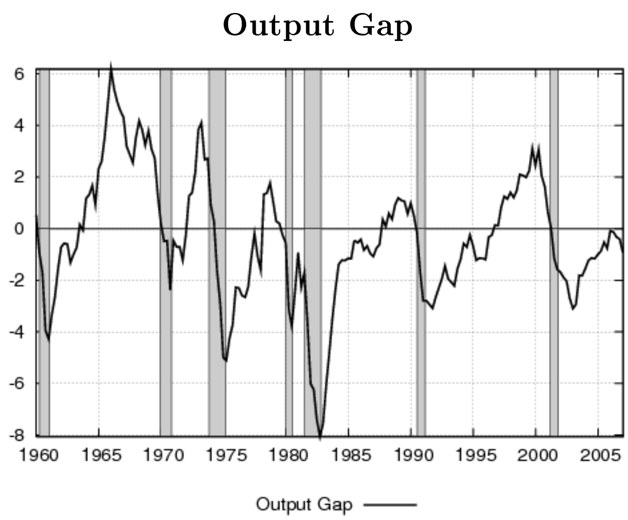
STOCK, J. H., AND M. W. WATSON (2003): "Has the Business Cycle Changed? Evidence and Expectations," Monetary Policy and Uncertainty, Federal Reserve Bank of Kansas City.

TAYLOR, J. (1993): "Discretionary Versus Policy Rules in Practice," Carnegie-Rochester Conference Series on Public Policy, 39, 195–214.

WILLIAMS, N. (2005): "Adaptive Learning and Business Cycles," Working paper.

WOODFORD, M. (2003): Interest and Prices. Princeton University Press.

Figure 3.1: Output Gap and Inflation



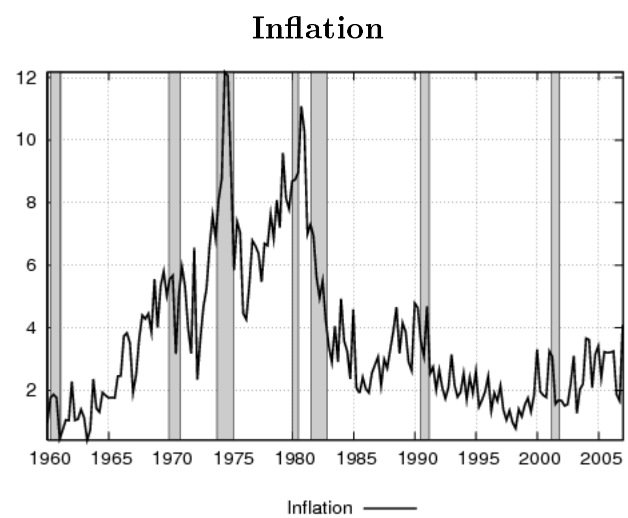


Table 3.1: Maximum Likelihood Parameter Estimates

		Rational Expectations		Dynamic Gain		Constant Gain	
Parameter	Description	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
η	Habit Formation	0.3643	0.0478	0.2580	0.0308	0.3659	0.0288
σ	Elasticity Substitution	0.0073	0.0154	0.2560	0.1171	0.1824	0.1140
μ	Elasticity Labor Supply	0.0000	40.9507	0.3219	2.2075	0.0001	5.0920
κ	Phillips Coefficient	0.0011	0.0186	0.0237	0.0256	0.0054	0.0146
γ	Price Indexation	0.8945	0.0330	0.9849	0.1926	0.9990	0.0004
ρ_r	MP Persistence	0.9355	0.0289	0.9234	0.0084	0.9196	0.0092
ψ_y	MP Output	0.2507	0.0498	0.1878	0.0367	0.2758	0.0425
ψ_π	MP Inflation	1.9577	0.2591	1.7363	0.1687	1.6354	0.1189
ρ_n	Nat. Rate Pers.	0.8705	0.0353	0.7484	0.0267	0.6936	0.0272
ρ_u	Cost Push Pers.	0.0000	0.0000	0.0062	0.0376	0.0031	0.0085
π_*	SS Inflation	3.5446	0.2808	4.4419	0.2220	5.3272	0.2825
$\sigma_{n,L}$	Nat. Rate (Low)	0.1768	0.3720	0.0454	0.0217	0.0931	0.0572
$\sigma_{u,L}$	Cost Push (Low)	0.0023	0.0001	0.0045	0.0004	0.0042	0.0001
$\sigma_{r,L}$	MP Shock (Low)	0.0013	0.0001	0.0012	0.0000	0.0012	0.0000
$\sigma_{n,H}$	Nat. Rate (High)	0.4295	0.9056	0.0966	0.0485	0.1794	0.1144
$\sigma_{u,H}$	Cost Push (High)	0.0044	0.0004	0.0092	0.0010	0.0085	0.0005
$\sigma_{r,H}$	MP Shock (High)	0.0070	0.0005	0.0064	0.0003	0.0056	0.0002
p_L	P(Remain Low)	0.9609	0.0224	0.9724	0.0097	0.9780	0.0109
p_H	P(Remain High)	0.8099	0.0578	0.8924	0.0264	0.9412	0.0159
g	Learning Gain	–	–	0.0045	0.0007	0.0000	0.0018

Table 3.2: Model Comparisons

	Rational Expectations	Dynamic Gain	Constant Gain
RMSE Output Gap	3.12	3.13	3.18
RMSE Inflation	4.41	4.69	4.69
RMSE Federal Funds Rate	5.01	5.05	5.09
AR(1) Output Variance	0.0904 (0.0730)	0.1715 (0.0722)	0.1240 (0.0728)
AR(1) Inflation Variance	0.1760 (0.0716)	0.1364 (0.0699)	0.1073 (0.0653)
AR(1) Fed Funds Variance	0.3851 (0.0670)	0.3798 (0.0659)	0.3798 (0.0636)

Figure 3.2: Smoothed Probability in Volatile State

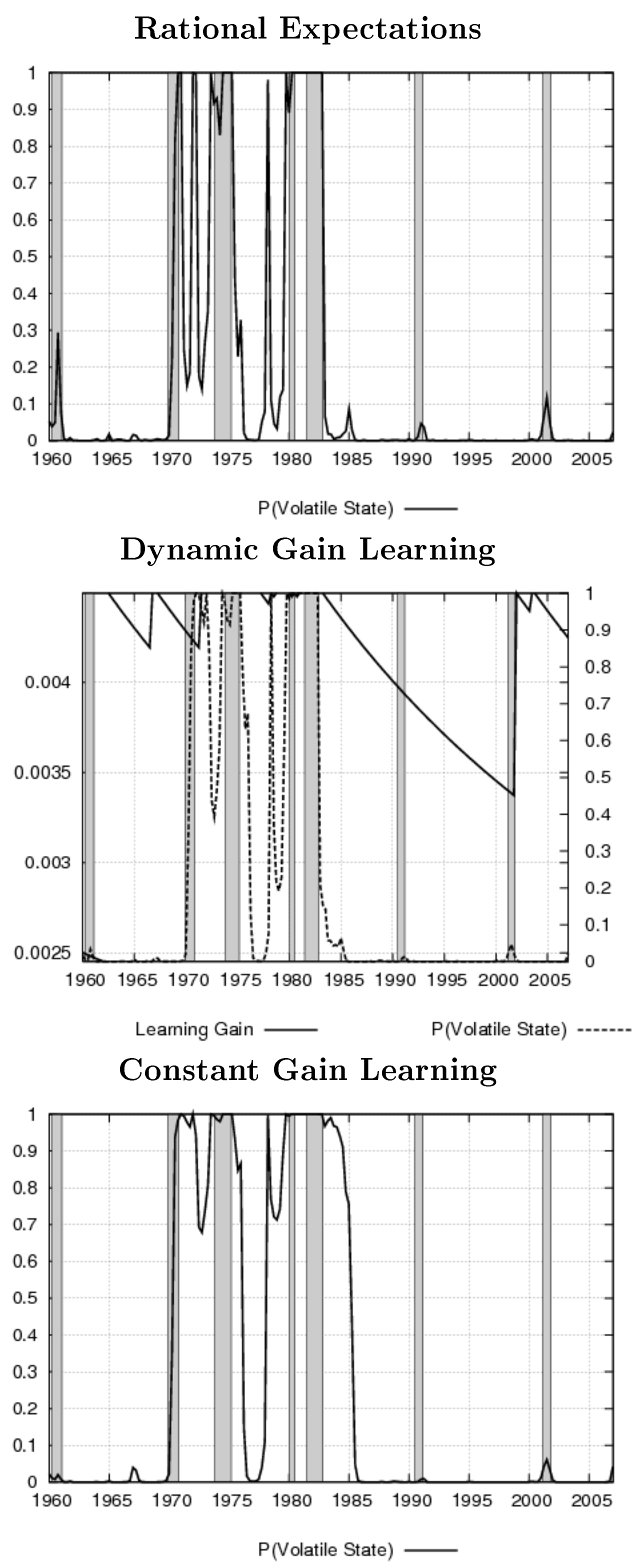


Figure 3.3: Smoothed Estimate of Natural Rate Shock

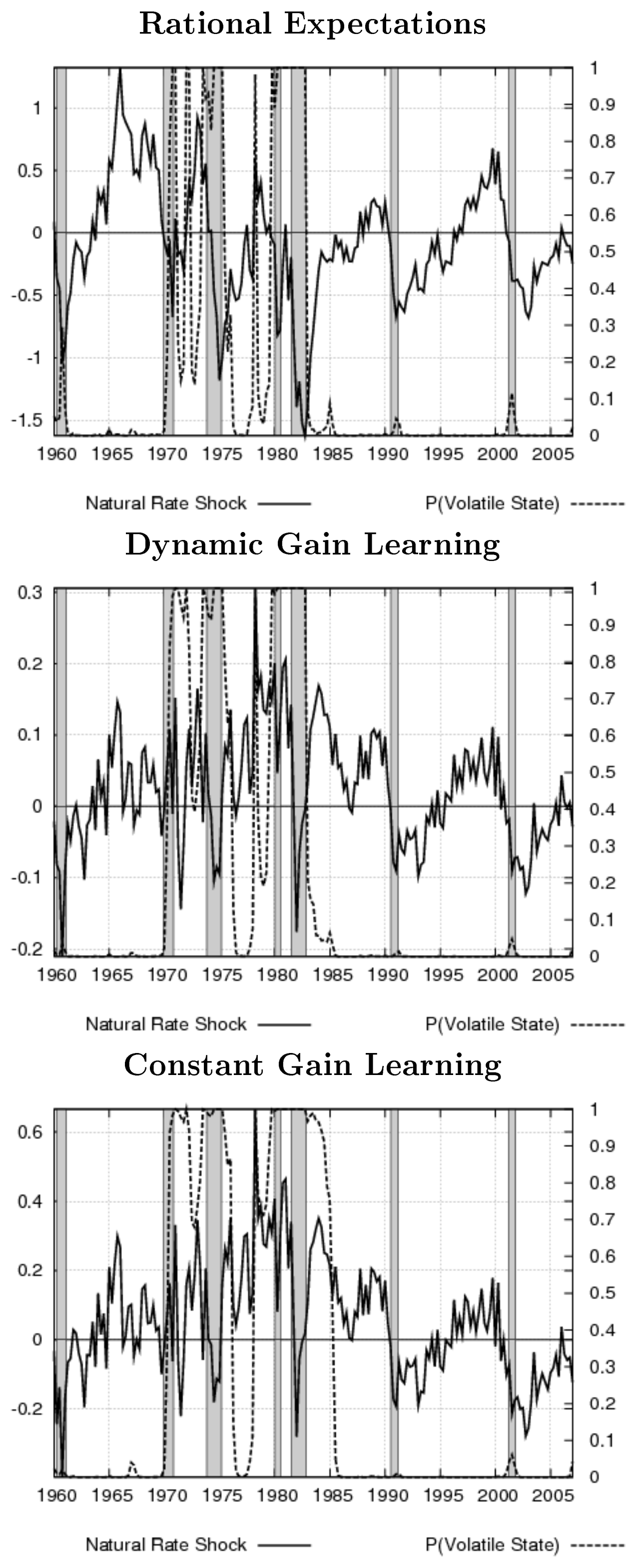


Figure 3.4: Smoothed Estimate of Cost Push Shock

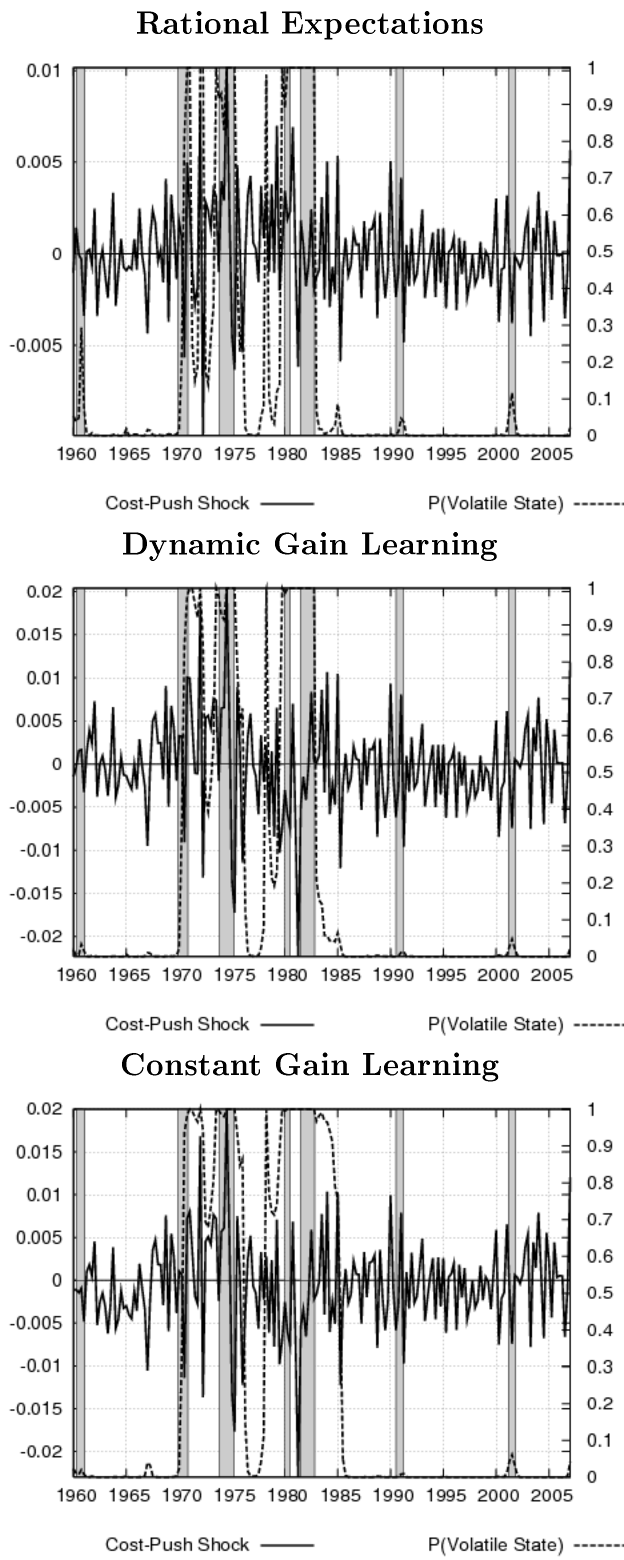


Figure 3.5: Smoothed Estimate of Monetary Policy Shock

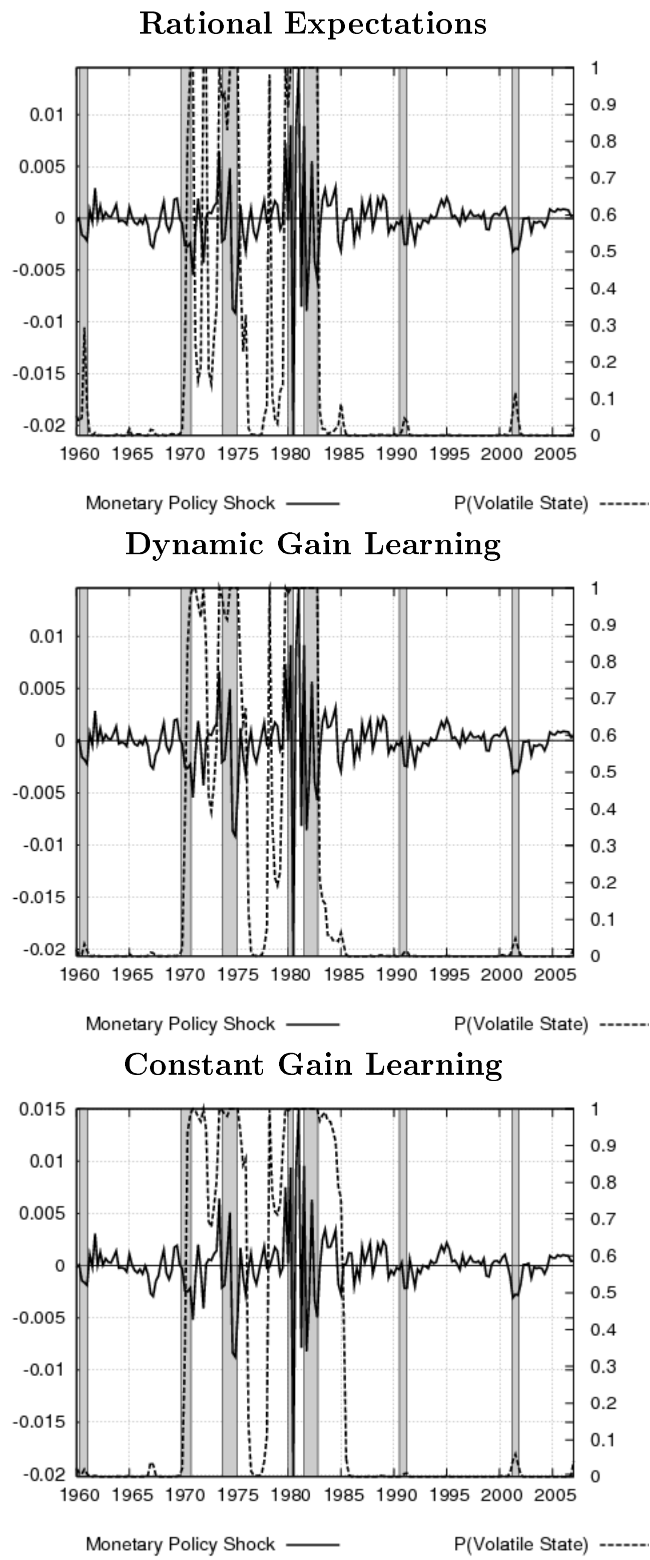
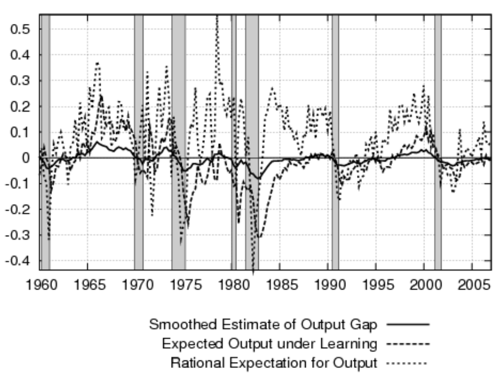


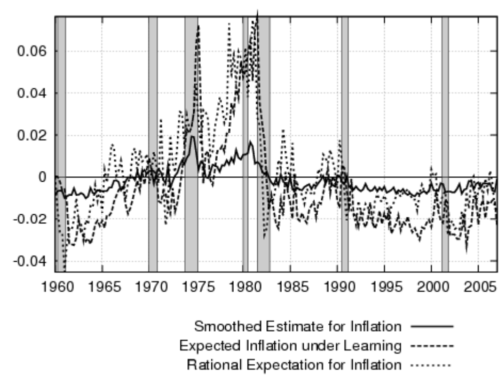
Figure 3.6: Agents' Expectations

Dynamic Gain Learning

Output Gap

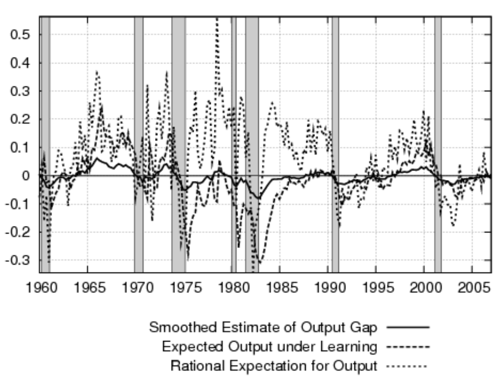


Inflation



Constant Gain Learning

Output Gap



Inflation

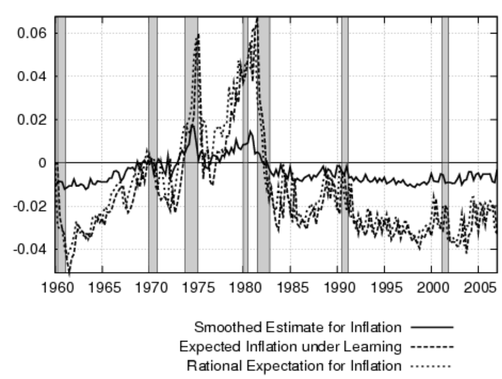
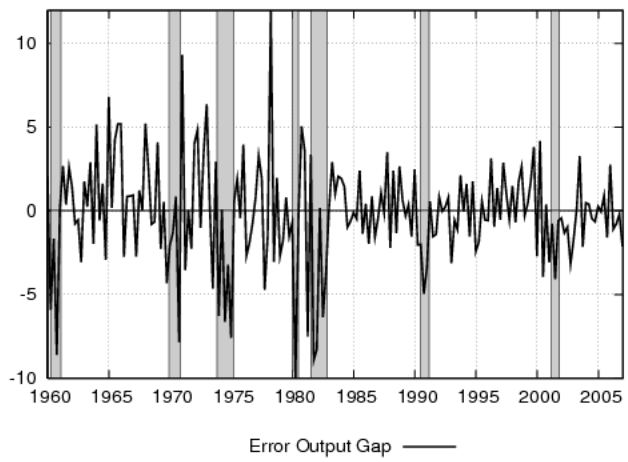
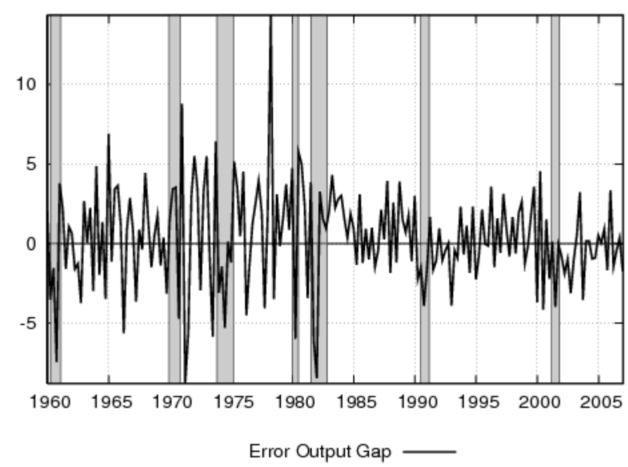


Figure 3.7: One Period Ahead Output Forecast Error

Rational Expectations



Dynamic Gain Learning



Constant Gain Learning

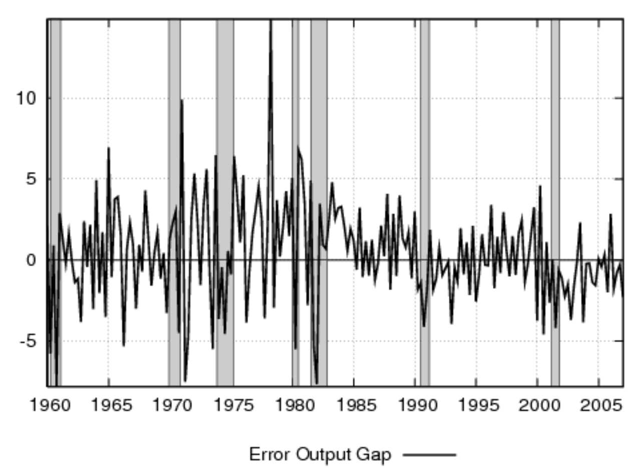


Figure 3.8: One Period Ahead Inflation Forecast Error

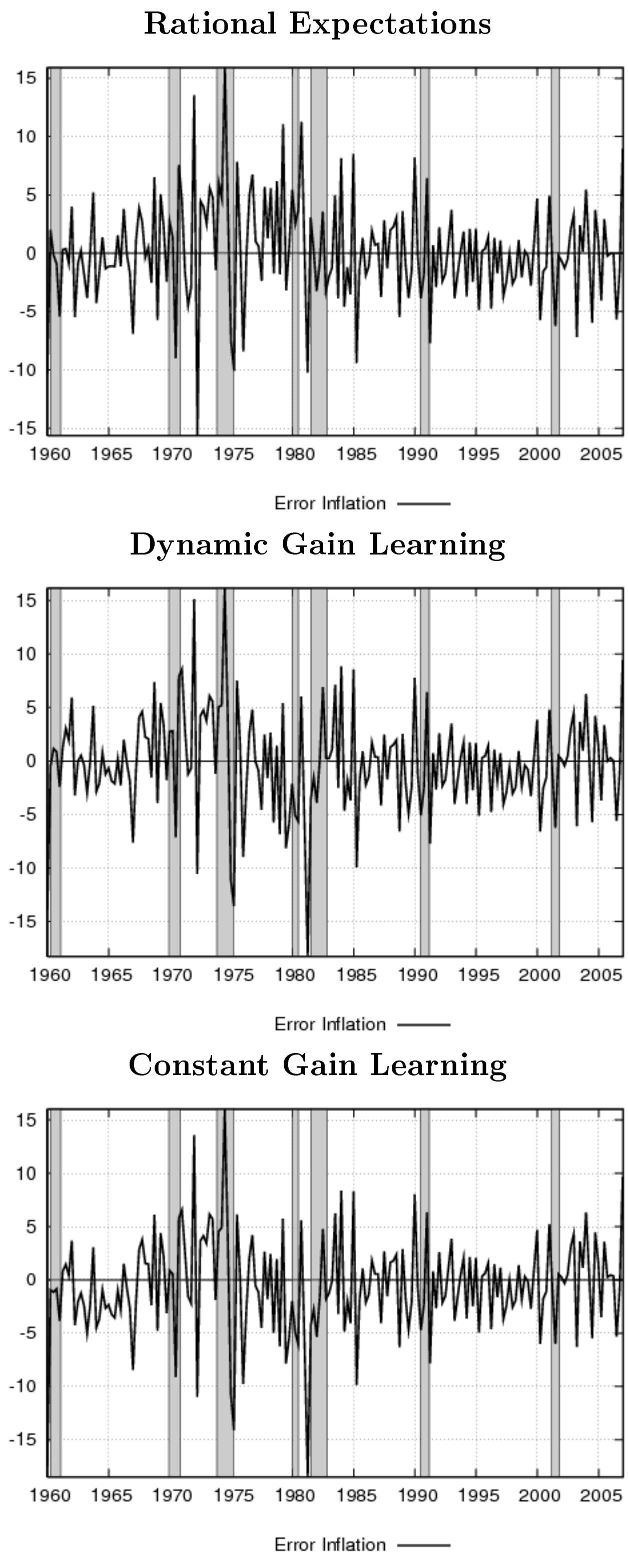
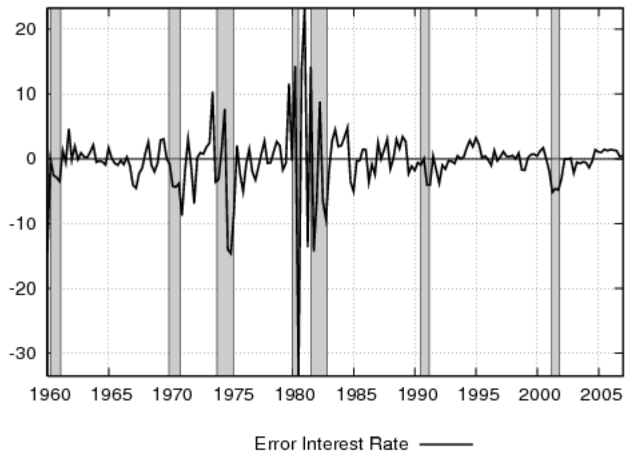
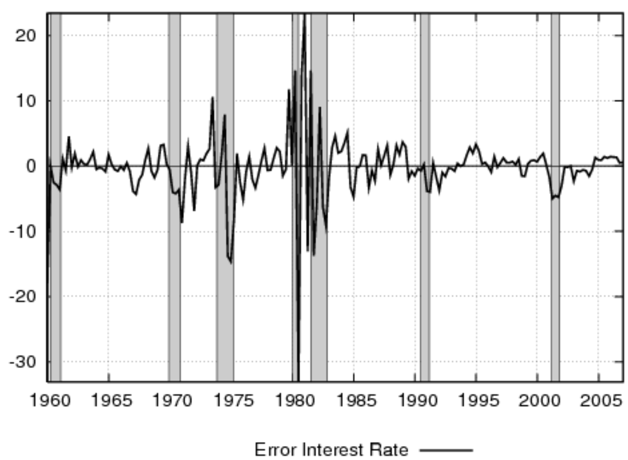


Figure 3.9: One Period Ahead Federal Funds Rate Forecast Error

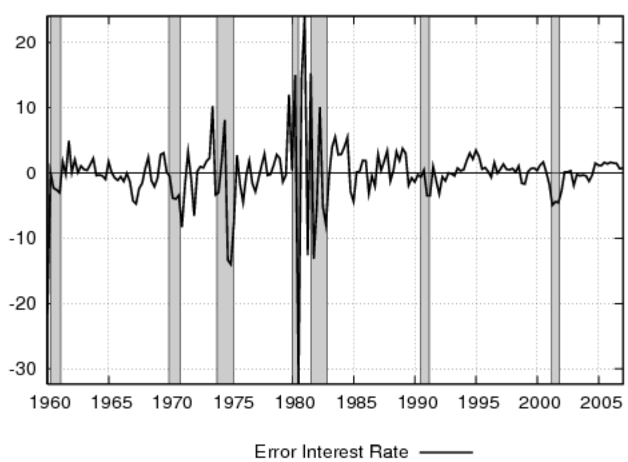
Rational Expectations



Dynamic Gain Learning



Constant Gain Learning



Appendix A

New Keynesian Model with Firm-Specific Capital: Derivations

A.1 Consumers

Consumers choose consumption, c_t, labor supply, $n_t(i)$, and purchases of bonds, $b_t(i)$, to maximize utility, given in equation (2.1), subject to the budget constraint, given in equation (2.2). The first order conditions are,

$$\lambda_t = \xi_t \left(c_t - \eta c_{t-1}\right)^{-\frac{1}{\sigma}} - \beta\eta E_t \xi_{t+1} \left(c_{t+1} - \eta c_t\right)^{-\frac{1}{\sigma}}$$

$$\mu_t n_t(i)^{\mu} = \lambda_t \frac{w_t(i)}{p_t}$$

$$\lambda_t = \beta E_t \lambda_{t+1} \frac{1+r_t}{1+\pi_{t+1}}$$

where λ_t is the Lagrange multiplier for the budget constraint and therefore the marginal utility of real income. Log-linearizing the first order conditions yields,

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

$$\hat{w}_t(i) - \hat{p}_t = \mu\hat{n}_t(i) - \hat{\lambda}_t + \hat{\mu}_t \quad \text{(A.2)}$$

$$\hat{\lambda}_t = E_t\hat{\lambda}_{t+1} + \hat{r}_t - E_t\pi_{t+1} \quad \text{(A.3)}$$

where a hat indicates the percentage deviation of the variable from its steady state. Equation (A.2) will be referenced later to express equilibrium real wages in terms of employment. Equations (A.3) and (A.1) together implicitly define the log-linear Euler equation which determines consumers' demand for final goods.

A.2 Producers

A.2.1 Final goods firms

The final goods firm chooses its demand for intermediate good i to maximize profits,

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

The first order condition leads to the demand for intermediate good i,

$$y_t(i) = \left[\frac{p_t(i)}{p_t} \right]^{-\theta} y_t. \tag{A.4}$$

which is given in equation (2.7).

A.2.2 Input choices

Intermediate goods firms choose labor demand and rent capital to minimize real total cost, given in equation (2.10), subject to the production function, given in equation (2.9). The first order conditions are,

$$\frac{w_t(i)}{p_t} = (1-\alpha) s_t(i) \frac{y_t(i)}{n_t(i)}, \tag{A.5}$$

$$\rho_t(i) = \alpha s_t(i) \frac{y_t(i)}{k_t(i)}, \tag{A.6}$$

where $s_t(i)$ is the Lagrange multiplier on the production function. The Lagrange multiplier is interpreted as the change in the objective function from a marginal ease in the constraint. In this case the objective function is total cost and the constraint is total output, so the Lagrange multiplier is equal to the marginal cost.

Log-linearizing the first order conditions yields,

$$\hat{\rho}_t(i) = \hat{s}_t(i) + \hat{y}_t(i) - \hat{k}_t(i), \tag{A.7}$$

$$\hat{w}_t(i) - \hat{p}_t = \hat{s}_t(i) + \hat{y}_t(i) - \hat{n}_t(i), \tag{A.8}$$

Combining these two equations to eliminate $\hat{s}_t(i)$ and substituting equation (A.2) to eliminate wages and prices leads to the expression for the rental rate of capital,

$$\hat{\rho}_t(i) = (\mu + 1)\,\hat{n}_t(i) - \hat{k}_t(i) - \hat{\lambda}_t + \hat{\mu}_t. \tag{A.9}$$

The production function can now be used to express the rental rate of capital only in terms of output and capital. The log-linear production function is given by,

$$\hat{y}_t(i) = \hat{z}_t + \alpha\hat{k}_t(i) + (1-\alpha)\hat{n}_t(i). \tag{A.10}$$

Solving equation (A.10) for $\hat{n}_t(i)$ and substituting this into (A.9) yields,

$$\hat{\rho}_t(i) = \frac{\mu+1}{1-\alpha}\hat{y}_t(i) - \frac{1+\mu\alpha}{1-\alpha}\hat{k}_t(i) - \hat{\lambda}_t + \hat{\mu}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t \tag{A.11}$$

Solving equation (A.7) for $\hat{s}_t(i)$ and using equation (A.11) to substitute out $\hat{\rho}_t(i)$ leads to the expression for marginal cost for firm i,

$$\hat{s}_t(i) = \frac{\alpha+\mu}{1-\alpha}\hat{y}_t(i) - \frac{1+\alpha\mu}{1-\alpha}\hat{k}_t(i) - \hat{\lambda}_t + \hat{\mu}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t \tag{A.12}$$

Summing over all the firms leads to the average marginal cost in the economy,

$$\hat{s}_t = \frac{\alpha+\mu}{1-\alpha}\hat{y}_t - \frac{\alpha(\mu+1)}{1-\alpha}\hat{k}_t - \hat{\lambda}_t + \hat{\mu}_t - \frac{\mu+1}{1-\alpha}\hat{z}_t \tag{A.13}$$

Subtracting equation (A.13) from equation (A.12), leads to an expression for the marginal cost of firm i in terms of the average marginal cost and the firms relative output and capital stock,

$$\hat{s}_t(i) = \hat{s}_t + \frac{\alpha + \mu}{1 - \alpha} \left[\hat{y}_t(i) - \hat{y}_t \right] - \frac{\alpha + \mu}{1 - \alpha} \tilde{k}_t(i) \tag{A.14}$$

where $\tilde{k}_t(i) = \hat{k}_t(i) - \hat{k}_t$ is the relative capital stock of firm i.

A.2.3 Capital goods firms

Capital goods firms maximize the utility value of profits, given in equation (2.16), subject to the evolution of firm-specific capital stock, given in equation (2.14). Instead of explicitly computing the profit maximizing choice of investment, one can solve the evolution of capital for $I_t(i)$ and substitute this into the objective function. The first order condition is,

$$\begin{gathered}
\frac{\lambda_t}{\iota_t} \left[1 + \phi \left(\frac{k_{t+1}(i)}{k_t(i)} - 1 \right) \right] = \\
\beta E_t \frac{\lambda_{t+1}}{\iota_{t+1}} \left[\iota_{t+1} \rho_{t+1}(i) + (1 - \delta) + \phi \left(\frac{k_{t+2}(i)}{k_{t+1}(i)} - 1 \right) \frac{k_{t+2}(i)}{k_{t+1}(i)} - \frac{\phi}{2} \left(\frac{k_{t+2}(i)}{k_{t+1}(i)} - 1 \right)^2 \right].
\end{gathered} \tag{A.15}$$

Log-linearizing this yields,

$$\begin{gathered}
\hat{\lambda}_t + \phi \left(\hat{k}_{t+1}(i) - \hat{k}_t(i) \right) = E_t \hat{\lambda}_{t+1} + \left[1 - \beta \left(1 - \delta \right) \right] E_t \hat{\rho}_{t+1}(i) \\
+ \beta \phi \left(E_t \hat{k}_{t+2}(i) - \hat{k}_{t+1}(i) \right) + \hat{\iota}_t - \beta (1 - \delta) E_t \hat{\iota}_{t+1}.
\end{gathered} \tag{A.16}$$

Plugging equation (A.11) into (A.16) leads to the following equilibrium condition for the evolution of capital stock for firm i:

$$\hat{\lambda}_t + \phi\left(\hat{k}_{t+1}(i) - \hat{k}_t(i)\right) = \beta(1-\delta)E_t\hat{\lambda}_{t+1}$$

$$+\left(\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right)\left[(\mu+1)E_t\hat{y}_{t+1}(i) - (1+\mu\alpha)\hat{k}_{t+1}(i)\right] + \beta\phi\left(E_t\hat{k}_{t+2}(i) - \hat{k}_{t+1}(i)\right)$$

$$+\left[1-\beta(1-\delta)\right]E_t\hat{\mu}_{t+1} - \frac{(\mu+1)\left[1-\beta\left(1-\delta\right)\right]}{1-\alpha}E_t\hat{z}_{t+1} + \hat{\iota}_t - \beta(1-\delta)E_t\hat{\iota}_{t+1}. \tag{A.17}$$

Integrating equation (A.17) over all firms leads to the evolution of the aggregate capital stock,

$$\hat{\lambda}_t + \phi\left(\hat{k}_{t+1} - \hat{k}_t\right) = \beta(1-\delta)E_t\hat{\lambda}_{t+1} + \left(\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right)\left[(\mu+1)E_t\hat{y}_{t+1} - (1+\mu\alpha)\hat{k}_{t+1}\right]$$

$$+\beta\phi\left(E_t\hat{k}_{t+2} - \hat{k}_{t+1}\right) - \frac{(\mu+1)\left[1-\beta\left(1-\delta\right)\right]}{1-\alpha}E_t\hat{z}_{t+1} + \hat{\iota}_t - \beta(1-\delta)E_t\hat{\iota}_{t+1}$$
$$+\left[1-\beta(1-\delta)\right]E_t\hat{\mu}_{t+1}, \tag{A.18}$$

which is shown in equation (2.17) of the paper. Subtracting equation (A.18) from equation (A.17) leads to following expression for firm i's capital stock in terms of the aggregate capital stock,

$$\begin{aligned}\phi\left(\tilde{k}_{t+1}(i) - \tilde{k}_t(i)\right) &= \beta\phi\left(E_t\tilde{k}_{t+2}(i) - \tilde{k}_{t+1}(i)\right)\\ &+\left[\frac{1-\beta\left(1-\delta\right)}{1-\alpha}\right]\left[(\mu+1)E_t\left(\hat{y}_{t+1}(i) - \hat{y}_{t+1}\right) - (1+\mu\alpha)\tilde{k}_{t+1}(i)\right].\end{aligned} \tag{A.19}$$

A.2.4 Optimal pricing

The inflation indexation rule given in equation (2.20) can be re-written so that future prices intermediate goods firms will charge while not being able to re-optimize their price can be expressed in terms of the price chosen by the firm at time t. By repeated substitution of equation (2.20), the price at time $t+T$ of good i can be expressed as,

$$p_{t+T}(i) = p_t(i) \exp\left(\gamma \sum_{\tau=0}^{T-1} \pi_{t+\tau}\right),$$

For notational convenience, let $\pi^*_{t+T} \equiv \sum_{\tau=0}^{T-1} \pi_{t+\tau}$. Substitute the demand equation, (2.7), into the profit function, (2.19), to express the profit only in terms of the intermediate good price, $p_t(i)$, and aggregate state variables the firm cannot control:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left(\frac{p_t(i) e^{\gamma \pi^*_{t+T}}}{p_{t+T}}\right)^{1-\theta} y_{t+T} - S\left[\left(\frac{p_t(i) e^{\gamma \pi^*_{t+T}}}{p_{t+T}}\right)^{-\theta} y_{t+T}\right]\right\}. \tag{A.20}$$

The first order condition with respect to $p_t(i)$ is given by,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ (1-\theta)\left(\frac{p^*_t(i) e^{\gamma \pi^*_{t+T}}}{p_{t+T}}\right)^{1-\theta} + \theta s_{t+T}(i) \left(\frac{p^*_t(i) e^{\gamma \pi^*_{t+T}}}{p_{t+T}}\right)^{-\theta}\right\} \frac{y_{t+T}}{p^*_t(i)} = 0, \tag{A.21}$$

where $p^*_t(i)$ is the optimal price for a firm that is able to re-optimize its price. Since the first order condition cannot be rewritten in terms of inflation instead of prices, it is necessary to assume prices have a steady state, which implies the steady state level of inflation is equal to zero. Before log-linearizing, it is convenient to rearrange equation (A.21) as,

$$\begin{aligned} (1-\theta) E_t \sum_{T=0}^{\infty} (\omega\beta)^T \lambda_{t+T} \left(\frac{p^*_t(i) e^{\gamma \pi^*_{t+T}}}{p_{t+T}}\right)^{1-\theta} y_{t+T} = \\ -\theta E_t \sum_{T=0}^{\infty} (\omega\beta)^T \lambda_{t+T} s_{t+T}(i) \left(\frac{p^*_t(i) e^{\gamma \pi^*_{t+T}}}{p_{t+T}}\right)^{-\theta} y_{t+T}, \end{aligned} \tag{A.22}$$

then log-linearize each side of the equal sign separately. Log-linearizing the left hand side and right hand size, respectively, yield,

$$(1-\theta)\lambda y E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{\lambda}_{t+T} + \hat{y}_{t+T} + (1-\theta)\left(\hat{p}^*_t(i) - \hat{p}_{t+T} + \gamma \pi^*_{t+T}\right)\right], \tag{A.23}$$

$$-\theta\lambda ys E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{\lambda}_{t+T} + \hat{y}_{t+T} + \hat{s}_{t+T} - \theta\left(\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^*\right)\right] \tag{A.24}$$

where λ is the steady state marginal utility of income, y is the steady state level of output, and s is the steady state marginal cost. Steady state marginal utility and steady output cancel out from the left and right hand sides. The steady state marginal cost is found by evaluating the first order condition (A.21) where $\lambda_t = \lambda$ and $p_t^*(i) = p_t = p$ for all t. In the steady state equation (A.21) simplifies to,

$$\left(\frac{1}{1-\omega\beta}\right)\frac{(1-\theta+\theta s)\,y}{p} = 0.$$

The steady state solution for s is given by,

$$s = -\frac{1-\theta}{\theta}. \tag{A.25}$$

The coefficient $-\theta\ s$ in equation (A.24) therefore cancels out with $1-\theta$ in equation (A.23). Combining the left and right hand side then yields,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* - \hat{s}_{t+T}(i)\right] = 0 \tag{A.26}$$

Solving for $\hat{p}_t^*(i)$ yields,

$$\hat{p}_t^*(i) = (1-\omega\beta)\,E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \hat{s}_{t+T}(i)\right]. \tag{A.27}$$

Substitute into equation (A.27), the log-linearized the demand for intermediate good i at time $t+T$, which is given by,

$$\hat{y}_{t+T}(i) = -\theta(\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^*) + \hat{y}_{t+T} \tag{A.28}$$

and the marginal cost given in equation (A.14). This leads to an expression for the optimal price for firm i in terms of aggregate variables and the firm's expected future capital,

$$\hat{p}_t^*(i) = \quad (1-\omega\beta)\, E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left\{ \hat{p}_{t+T} - \gamma\pi_{t+T}^* + \hat{s}_{t+T} - \frac{\theta\,(\alpha+\mu)}{1-\alpha} \left[\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* \right] \right\}$$

$$- (1-\omega\beta)\, E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left\{ \frac{\alpha\,(\mu+1)}{1-\alpha} \tilde{k}_{t+T}(i) \right\}. \tag{A.29}$$

The solution of this equation for $\hat{p}_t^*(i)$ is given by,

$$\hat{p}_t^*(i) = (1-\omega\beta)\, E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \psi\hat{s}_{t+T} - \frac{\psi\alpha(\mu+1)}{1-\alpha} \tilde{k}_{t+T}(i) \right], \tag{A.30}$$

where

$$\psi = \left[1 + \frac{\theta(\alpha+\mu)}{1-\alpha} \right]^{-1}.$$

Equation (A.30) can be rewritten as the first order difference equation:

$$\hat{p}_t^*(i) = \omega\beta E_t \hat{p}_{t+1}^*(i) + (1-\omega\beta) \left(\hat{p}_t + \psi\hat{s}_t - \frac{\psi\alpha(\mu+1)}{\mu(1-\alpha)} \tilde{k}_t(i) \right), \tag{A.31}$$

where $E_t\hat{p}_{t+1}^*(i)$ denotes the expectation at time t for the time $t+1$ optimal decision for the firm's new price, conditional that the firm is able to re-optimize its price again in period $t+1$. Note, this is not the same as the unconditional time t expectation of the firm's price in period $t+1$. Since with probability ω the firm will not be able to re-optimize its price next period, the unconditional expectation for firm i's price in period $t+1$ is given by,

$$E_t\hat{p}_{t+1}(i) = \omega \left[\hat{p}_t^*(i) + \gamma\pi_{t-1} \right] + (1-\omega) E_t \hat{p}_{t+1}^*(i). \tag{A.32}$$

A.2.5 Phillips Curve Solution

Deriving the Phillips curve when there is firm-specific capital is substantially more complicated than a model without capital or with a perfect capital rental market. Equation (A.31) shows that each firm's optimal price will depend on its capital stock relative to the aggregate capital stock. Since a firm's capital stock is dependent on its entire investment history, the optimal price will depend on the firm's entire history of being able to re-optimize its price. The convenient result from the previous section that each firm will choose the same price does not hold when there is firm-specific capital and Calvo pricing.

Equation (A.29) implicitly defines the optimal choice for the price of intermediate good i in terms of expectations of aggregate variables and the following expectation of the firm's future relative capital stocks:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) \tag{A.33}$$

To derive the Phillips curve, we must rewrite the above expression in terms of the firm's current capital stock, the current optimal price, and expectations of aggregate variables. The optimal choice for $\tilde{k}_{t+1}(i)$ in terms of expected future output is given in equation (A.19). Substituting the log-linear demand for $y_{t+1}(i)$ into equation (A.19) leads to,

$$\begin{aligned}\phi\left(\tilde{k}_{t+1}(i) - \tilde{k}_t(i)\right) &= \beta\phi\left(E_t\tilde{k}_{t+2}(i) - \tilde{k}_{t+1}(i)\right) \\ &- \left[\frac{1-\beta(1-\delta)}{1-\alpha}\right]\left[\theta(\mu+1)E_t\tilde{p}_{t+1}(i) - (1+\mu\alpha)\tilde{k}_{t+1}(i)\right].\end{aligned} \tag{A.34}$$

where $\tilde{p}_{t+1}(i) = \hat{p}_{t+1}(i) - \hat{p}_{t+1}$ is the relative price of intermediate good i in period $t+1$. The rational expectations solution for (A.34) must have the form,

$$\tilde{k}_{t+1}(i) = m\tilde{k}_t(i) + n\tilde{p}_t(i), \tag{A.35}$$

where m and n are determined by the method of undetermined coefficients in the next subsection. For a firm re-optimizing their price, this equation can be rewritten as

$$\tilde{k}_{t+1}(i) = m\tilde{k}_t(i) + n\hat{p}_t^*(i) - n\hat{p}_t(i). \tag{A.36}$$

Substituting this into equation (A.33) shows that,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T+1}(i) = mE_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) + \frac{n}{1-\omega\beta}\hat{p}_t^*(i) - nE_t \sum_{T=0}^{\infty} (\omega\beta)^T \hat{p}_{t+T}.$$

Multiply both sides of this equation by $(\omega\beta)$ then add $\tilde{k}_t(i)$ to both sides in order to make the summation on the left hand side identical to the summation on the right hand side. Doing this yields,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) = \omega\beta m E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) + \frac{\omega\beta n}{1-\omega\beta}\hat{p}_t^*(i) - \omega\beta n E_t \sum_{T=0}^{\infty} (\omega\beta)^T \hat{p}_{t+T} + \tilde{k}_t(i).$$

Solving this equation yields,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) = \frac{1}{1-\omega\beta m}\left[\frac{\omega\beta n}{1-\omega\beta}\hat{p}_t^*(i) + \tilde{k}_t(i) - \omega\beta n E_t \sum_{T=0}^{\infty} (\omega\beta)^T \hat{p}_{t+T}\right]. \tag{A.37}$$

Substituting this into equation (A.29) and solving for $\hat{p}_t^*(i)$ leads to the following solution,

$$\hat{p}_t^*(i) = (1-\omega\beta)E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left(\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \nu\hat{s}_{t+T}\right) - \frac{\alpha\nu(\mu+1)(1-\omega\beta)}{(1-\alpha)(1-\omega\beta m)}\tilde{k}_t(i), \tag{A.38}$$

where,

$$\nu = \left[1 + \frac{\theta(\alpha+\mu)}{1-\alpha} + \frac{\alpha\omega\beta n(\mu+1)}{(1-\alpha)(1-\omega\beta m)}\right].$$

Equation (A.38) expresses the optimal price of intermediate good i solely in terms of aggregate variables and the firm's current relative capital stock. Since the capital stock was chosen in the previous period, it is independent of whether or not a firm is currently able to re-optimize its price. Therefore the average capital stock among firms re-optimizing their price is equal to the average capital stock in the economy. This implies that average value for $\tilde{k}_t(i)$ over firms re-optimizing their price is equal to zero. Let $\hat{p}_t^*$ denote the average price among these firms. Equation (A.38) implies,

$$\hat{p}_t^* = (1-\omega\beta)E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left(\hat{p}_{t+T} - \gamma\pi_{t+T}^* + \nu\hat{s}_{t+T}\right). \tag{A.39}$$

This can be rewritten as the first order difference equation,

$$\hat{p}_t^* = \omega\beta E_t \hat{p}_{t+1}^* + (1-\omega\beta)\left(\hat{p}_t + \nu\hat{s}_t\right). \tag{A.40}$$

Substituting equation (A.39) into (A.40) to eliminate $\hat{p}_t^*$ and $E_t\hat{p}_{t+1}^*$ leads to the Phillips curve,

$$\pi_t = \left(\frac{1}{1+\beta\gamma}\right)\left[\gamma\pi_{t-1} + \beta E_t\pi_{t+1} + \kappa\hat{s}_t\right], \tag{A.41}$$

where,

$$\kappa = \frac{(1-\omega)(1-\omega\beta)}{\nu\omega}.$$

A.2.6 Method of Undetermined Coefficients

This subsection uses the method of undetermined coefficients to compute the values of m and n in equation (A.36) which must satisfy the optimality condition for capital given in equation (A.34). Equation (A.34) can be rearranged as,

$$\tilde{k}_{t+1}(i) = \tilde{k}_t(i) + \beta E_t\tilde{k}_{t+2}(i) - \zeta_0 E_t\tilde{p}_{t+1}(i) - \zeta_1\tilde{k}_{t+1}(i), \tag{A.42}$$

where ζ_0 and ζ_1 are given by,

$$\zeta_0 = \frac{\theta\left(\mu+1\right)\left[1-\beta\left(1-\delta\right)\right]}{\phi\left(1-\alpha\right)}$$

$$\zeta_1 = \beta + \frac{\left(1+\alpha\mu\right)\left[1-\beta\left(1-\delta\right)\right]}{\phi\left(1-\alpha\right)}$$

I begin by finding an expression for $E_t\tilde{p}_{t+1}(i)$ in terms of $\tilde{k}_t(i)$ and $\tilde{p}_t(i)$. Using equation (A.32), the expected relative price can be rewritten as,

$$E_t\tilde{p}_{t+1}(i) = E_t\hat{p}_{t+1}(i) - E_t\hat{p}_{t+1} = \omega\hat{p}_t(i) + (1-\omega)E_t\hat{p}_{t+1}^*(i) - E_t\hat{p}_{t+1} \tag{A.43}$$

In order to express $\tilde{p}_{t+1}(i)$ only in terms of $\tilde{p}_t(i)$ and $\tilde{k}_t(i)$, we must next find a solution for $\hat{p}_t^*(i)$. According to equation (A.31), the rational expectation solution for $\hat{p}_t^*(i)$

must take the form,

$$\hat{p}_t^*(i) = f(\hat{p}_t, \hat{s}_t) + a\tilde{k}_t(i), \tag{A.44}$$

where $f(\cdot)$ is a linear function of aggregate variables and a needs to be determined by the method of undetermined coefficients. Let $\mathcal{L}_t$ denote the set of firms re-optimizing their price in period t. The average price of the firms who are able to re-optimize their price is given by,

$$\hat{p}_t^* = \frac{1}{1-\omega}\int_{i\in\mathcal{L}_t} \hat{p}_t^*(i)di = f(\hat{p}_t, \hat{s}_t) + \frac{a}{1-\omega}\int_{i\in\mathcal{L}_t} \tilde{k}_t(i)di$$

Since $\tilde{k}_t(i)$ was chosen in period $t-1$, it is independent of whether a firm is re-optimizing its price. Therefore the average difference between a firm's capital stock and the aggregate capital stock among firms re-optimizing their price is equal to zero. Therefore,

$$\hat{p}_t^* = f(\hat{p}_t, \hat{s}_t),$$

and equation (A.44) can be rewritten as,

$$\hat{p}_t^*(i) = \hat{p}_t^* + a\tilde{k}_t(i). \tag{A.45}$$

Advancing equation (A.45) one period and taking expectations yields,

$$E_t\hat{p}_{t+1}^*(i) = E_t\hat{p}_{t+1}^* + a\tilde{k}_{t+1}(i), \tag{A.46}$$

where $E_t\hat{p}_{t+1}^*$ is the expected average price over firms that can re-optimize their price next period. This can be rewritten in terms of the expected aggregate price level. Since a fraction ω firms will not be able to change their price next period and the remaining $1-\omega$ firms will have an average price $\hat{p}_{t+1}^*$, the expected price level next period is given by,

$$E_t\hat{p}_{t+1} = \omega\hat{p}_t + (1-\omega)E_t\hat{p}_{t+1}^*. \tag{A.47}$$

Solving (A.47) for $E_t\hat{p}^*_{t+1}$ and substituting this expression into (A.46) leads to,

$$E_t\hat{p}^*_{t+1}(i) = \frac{1}{1-\omega}\left(E_t\hat{p}_{t+1} - \omega\hat{p}_t\right) + a\tilde{k}_{t+1}(i). \tag{A.48}$$

Substituting equation (A.36) for $\tilde{k}_{t+1}(i)$ yields,

$$E_t\hat{p}^*_{t+1}(i) = \frac{1}{1-\omega}\left(E_t\hat{p}_{t+1} - \omega\hat{p}_t\right) + am\tilde{k}_t(i) - an\tilde{p}_t(i). \tag{A.49}$$

Plugging this into equation (A.43) leads to an expression for $E_t\tilde{p}_{t+1}(i)$ in terms of $\tilde{p}_t(i)$ and $\tilde{k}_t(i)$,

$$E_t\tilde{p}_{t+1}(i) = \left[\omega + (1-\omega)\,an\right]\tilde{p}_t(i) + (1-\omega)\,am\tilde{k}_t(i) \tag{A.50}$$

Next, using equation (A.36), the expected future capital stock is given by,

$$E_t\tilde{k}_{t+2}(i) = m^2\tilde{k}_t(i) + mn\tilde{p}_t(i) + nE_t\tilde{p}_{t+1}(i) \tag{A.51}$$

Substituting equation (A.50) into equation (A.51) leads to an expression for $E_t\tilde{k}_{t+2}(i)$ in terms of $\tilde{p}_t(i)$ and $\tilde{k}_t(i)$,

$$E_t\tilde{k}_{t+2}(i) = \Big[m^2 + amn\,(1-\omega)\Big]\,\tilde{k}_t(i) + \Big[mn + n\omega + an^2(1-\omega)\Big]\,\tilde{p}_t(i) \tag{A.52}$$

Plugging in equations (A.50), (A.51), and (A.36) into (A.42) leads to an expression for capital of the form given in equation (A.36) where m and n must satisfy, respectively,

$$\beta m^2 + \left[\beta an(1-\omega) - \zeta_1 - \zeta_0 a(1-\omega) - 1\right]m + 1 = 0, \tag{A.53}$$

$$\beta a(1-\omega)n^2 + \left[\beta m + \beta\omega - \zeta_1 - \zeta_0 a(1-\omega) - 1\right]n - \zeta_0\omega = 0. \tag{A.54}$$

All that remains is to find an expression for a, also using the method of undetermined coefficients. Substituting the expression for $E_t\hat{p}^*_{t+1}(i)$ given in equation (A.49) into equation (A.31) and solving for $\hat{p}_t^*(i)$ yields,

$$\begin{aligned}\hat{p}_t^*(i) = & \frac{1}{1-\omega\beta an}\left(\omega\beta am - \frac{\psi\alpha\,(\mu+1)}{(1-\alpha)}\right)\tilde{k}_t(i) \\ & + \frac{\omega\beta}{(1-\omega)\,(1-\omega\beta an)}\left(E_t\hat{p}_{t+1} - \omega\hat{p}_t\right) + \hat{p}_t + \frac{\psi}{1-\omega\beta an}\hat{s}_t,\end{aligned} \tag{A.55}$$

which implies a must satisfy the quadratic equation,

$$\omega\beta na^2 + (\omega\beta m - 1)\,a - \frac{\alpha\psi(\mu+1)}{1-\alpha} = 0. \tag{A.56}$$

Equations (A.53), (A.54), and (A.56) make up a system of quadratic equations that jointly determine the values for m, n, and a in terms of the parameters of the model. Since this is a system of three quadratic equations, there are potentially eight solutions, but these equations alone do not rule out economically infeasible outcomes. Equations (A.35) and (A.50) can be rewritten as the following dynamic system:

$$\begin{bmatrix} \tilde{k}_{t+1}(i) \\ E_t\tilde{p}_{t+1}(i) \end{bmatrix} = \begin{bmatrix} m & n \\ \omega + (1-\omega)an & 1-\omega \end{bmatrix} \begin{bmatrix} \tilde{k}_t(i) \\ \tilde{p}_t(i) \end{bmatrix}. \tag{A.57}$$

The economically feasible solution for m, n, and a must be consistent with stable means and variances of each firm's relative capital stock and relative price. The system is stable if and only if the eigenvalues of the matrix in equation (A.57) are inside the unit circle. The eigenvalues are given by,

$$e_1 = \frac{1}{2}\left(m + \omega + (1-\omega)an + \sqrt{\left[m + \omega + (1-\omega)an\right]^2 - 4m\omega}\right)$$

$$e_2 = \frac{1}{2}\left(m + \omega + (1-\omega)an - \sqrt{\left[m + \omega + (1-\omega)an\right]^2 - 4m\omega}\right)$$

It is evident from these equations that $e_1 > e_2$. Therefore both eigenvalues will be less than 1 in absolute value if and only if $e_1 < 1$ and $e_2 > -1$. The condition on the first eigenvalue implies,

$$\sqrt{[m+\omega+(1-\omega)an]^2-4m\omega}<2-m-\omega-(1-\omega)an. \tag{A.58}$$

Since the left hand side of the inequality is always positive, the left hand side must also be positive. Therefore, squaring both sides preserves the direction of the inequality. Doing this yields,

$$[m+\omega+(1-\omega)an]^2-4m\omega<4-4\left[m+\omega+(1-\omega)an\right]+4\left[m+\omega+(1-\omega)an\right]^2 \tag{A.59}$$

This inequality does not preserve the restriction implied in (A.58) that the right hand side be positive. Therefore (A.58) also implies

$$2-m-\omega-(1-\omega)an>0. \tag{A.60}$$

The inequalities (A.59) and (A.60) simplify to, respectively,

$$m<1-an \tag{A.61}$$

$$m<1+(1-\omega)(1-an) \tag{A.62}$$

The stability condition for the second eigenvalue is,

$$\sqrt{[m+\omega+(1-\omega)an]^2-4m\omega}<2+m+\omega+(1-\omega)an,$$

which simplifies to,

$$m>-1-\frac{1-\omega}{1+\omega}an \tag{A.63}$$

Finally, the coefficients m, n, and a can be found by the solving the system of quadratic equations (A.53), (A.54), and (A.56), subject to the inequalities (A.61), (A.62), and (A.63).

A.3 Market clearing

Goods market clearing implies total output of the final good is equal to aggregate consumption plus aggregate investment,

$$y_t = c_t + I_t.$$

Log-linearizing this yields,

$$\hat{y}_t = c_y \hat{c}_t + \delta k_y \hat{I}_t, \tag{A.64}$$

where c_y is the steady state consumption to output ratio and k_y is the steady state capital to output ratio. The steady state capital to output ratio is found by combining the steady state first order condition for capital rental, given in equation (A.6), and the steady state first order condition for investment, given in equation (A.15). Evaluating equation (A.6) at the steady state and using the steady state marginal cost, given in equation (A.25), yields,

$$\rho = \alpha \frac{\theta - 1}{\theta} \left(\frac{y}{k} \right).$$

Evaluating equation (A.15) at the steady state yields,

$$1 = \beta \left(\rho + 1 - \delta \right).$$

Combining these equations to eliminate ρ leads to the following capital to output ratio,

$$k_y = \frac{\beta\alpha(\theta - 1)}{\theta \left(1 - \beta + \beta\delta\right)} \tag{A.65}$$

Evaluating the goods market clearing condition, (A.64), at the steady state yields the following consumption to output ratio,

$$c_y = 1 - \delta k_y. \tag{A.66}$$

James Murray
Curriculum Vitae

Contact Information

Department of Economics
Indiana University
105 Wylie Hall
100 S. Woodlawn
Bloomington, IN 47405

Phone: (608) 738 5408
E-mail: jmmurray@indiana.edu
http://mypage.iu.edu/˜jmmurray/

Citizenship: United States

Education

Ph.D.	Economics,	Indiana University	September 2008
M.A.	Economics,	Indiana University	May 2004
M.A.	Economics,	University of Notre Dame	May 2002
B.S.	Economics,	University of Wisconsin - La Crosse	May 2000

Dissertation

"Three Essays in Adaptive Expectations in New Keynesian Monetary Economies." Advisor: Dr. Eric Leeper.

Working Papers

"Initial Expectations in New Keynesians Models with Learning"

"Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital"

"Regime Switching, Learning, and the Great Moderation"

"Estimating the Effects of Dormitory Living on Student Performance" with Pedro Falcão de Araujo.

Refereed Publications

"Shirking in Major League Baseball in the Era of the Reserve Clause." with Glenn Knowles, Michael Haupert, and Keith Sherony. *Nine: A Journal of Baseball History and Social Policy Perspectives.* Volume 9. Spring 2001.

Non-Refereed Publications

"Expectations for Monetary Policy." *Business Connection.* April 2008.

"Economic Outlook for Bio-Fuels." *Business Connection.* February 2008.

Research Interests

Learning and Expectations
Applied Macroeconometrics
Monetary Economics
Scholarship of Teaching and Learning

Conference and Seminar Presentations

Learning Week Conference, St. Louis Federal Reserve Bank, July 2007.
"Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital"

Indiana University Economics Department Brown Bag Workshop, May 2007.
"Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital"

Jordan River Conference, Indiana University, April 2007.
"Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital"

Jordan River Conference, Indiana University, April 2007.
Discussion of Allaby, "Feasibility of Corn Ethanol from a Land Use Perspective."

Missouri Economics Conference, University of Missouri, March 2007.
"Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital"

Indiana Academy of Social Sciences Annual Meeting, October 2006.
"Empirical Significance of Learning and the Consequences of Mis-specifying Expectations"

Jordan River Conference, Indiana University, April 2006.
"Empirical Significance of Learning and the Consequences of Mis-specifying Expectations"

Jordan River Conference, Indiana University, April 2005.
"Liquidity in a Two Country Open Economy Model: Evidence from United States and Germany"

Awards

Jordan River Conference Best Graduate Student Paper Award, April 2007.

Future Faculty Teaching Fellowship, 2007.

Teaching Interests

Undergraduate:

Principles and Intermediate Macroeconomics
Principles and Intermediate Microeconomics
Elementary Statistics
Statistical Methods / Econometrics
Monetary Economics
International Economics / Finance
Mathematical Economics
Operations Research

Graduate:

Statistical Methods / Research Methods
Applied Econometrics
Macroeconomics
Monetary Economics
Open Economy Macroeconomics
Computational Economics
Operations Research

Employment

Teaching Fellow	IUPU - Columbus	8/2007 - 5/2008
Adjunct Professor	Viterbo University	5/2007 - 8/2007
		5/2006 - 8/2006
Associate Instructor	Indiana University	9/2003 - 5/2007
Adjunct Professor	Ivy Tech State College	3/2004 - 8/2004
Teaching and Research Assistant	Indiana University	9/2002 - 5/2003
Teaching and Research Assistant	University of Notre Dame	9/2000 - 8/2002
Intern Computer Programmer	Trane Company	8/1999 - 8/2000

Teaching Experience

Primary Instructor:

Econ 202: Principles of Macroeconomics	IUPU - Columbus, 1 Semester
Econ 270: Introductory Statistics	IUPU - Columbus, 2 Semesters
Math 130: Introductory Statistics	Viterbo University, 2 Sessions
Econ E201: Principles of Microeconomics	Indiana University, 6 Semesters
Econ E202: Principles of Macroeconomics	Indiana University, 2 Semesters
Econ E322: Intermediate Macroeconomics	Indiana University, 2 Summers
Econ E201: Principles of Macroeconomics	Ivy Tech State College, 1 Session
Econ E202: Principles of Microeconomics	Ivy Tech State College, 1 Session

Teaching Assistant:

Econ E472: Econometrics II	Indiana University, 1 Semester
Econ S370: Honors Statistics	Indiana University, 1 Semester
Econ E471: Econometrics I	Indiana University, 1 Semester
Econ 201: Principles of Microeconomics	University of Notre Dame, 1 Semester
Econ 592: Graduate Econometrics I	University of Notre Dame, 1 Semester

Professional Development
AIMS (Adapting Innovative Materials for Statistics) Workshop, Minneapolis, MN. July 2008.

Service Learning Workshop hosted by Indiana University Purdue University - Indianapolis Center for Service and Learning. January 2008.

Indiana University FACET (Faculty Colloquium on Excellence in Teaching) Summer Institute.
July 2007.

Professional/Academic Service
Economics Candidate Search and Screen Committee. Indiana University Purdue University - Columbus. Fall 2007 - Spring 2008.

Community Service
Mentor for Big Brothers Big Sisters of Columbus, IN. January 2008 - May 2008.

Bartholomew Consolidated School Corporation (BCSC) Book Buddy. September 2007 - April 2008.

Science/Inventor Fair judge. Central Middle School, Columbus, IN. February 2008.

Computer Skills

Teaching Tools:	Blackboard, Oncourse CL, MyEconLab, CourseCompass, Discover Econ, E-instruction CPS (Classroom Performance System), Interwrite PRS (Personal Response System).
Econometrics / Statistics:	MatLab, Gauss, Octave, Maple, Stata, EViews, Limdep, RATS, SAS, SPSS, Statistica, GNU Scientific Libraries for C, LAPACK (Linear Algebra Package for Fortran).
Programming Languages:	C, C++, MS Visual C++, MS Visual Basic, Java, Fortran, Eiffel, Perl.
Web Programming:	Java, Perl, HTML, Javascript.
Operating Systems:	Linux, Unix, Windows.
Other:	LaTeX, Beamer for LaTeX, Emacs, SQL, Unix shell programming (BASH, CSH), MS Office.

References

Teaching References

Dr. Jan Eriksen
Dean
School of Adult Learning
Viterbo University
900 Viterbo Drive
La Crosse, WI 54601
Phone: (608) 796-3087
E-mail: `jperiksen@viterbo.edu`

Dr. James Self
Lecturer of Economics
Department of Economics
Indiana University
105 Wylie Hall
100 S. Woodlawn
Bloomington, IN 47405
Phone: (812) 855-7590
E-mail: `jkself@indiana.edu`

Research References

Dr. Eric Leeper
Professor of Economics
Department of Economics
Indiana University
105 Wylie Hall
100 S. Woodlawn
Bloomington, IN 47405
Phone: (812) 855-9157
E-mail: `eleeper@indiana.edu`

Dr. Kim Huynh
Assistant Professor of Economics
Department of Economics
Indiana University
105 Wylie Hall
100 S. Woodlawn
Bloomington, IN 47405
Phone: (812) 855-2288
E-mail: `kphuynh@indiana.edu`