

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

James Murray  
Dahl School of Business  
Viterbo University<sup>†</sup>

September 30, 2008

## 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.

*Keywords:* Learning, firm-specific capital, New Keynesian model, maximum likelihood.

*JEL classification:* C13, E22, E31, E50.

---

\*I am grateful for the advice and guidance of 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.

<sup>†</sup>*Mailing address:* 900 Viterbo Drive, La Crosse, WI 54601. *E-mail address:* jmmurray@viterbo.edu.  
*Phone number:* (608) 738-5408.

# 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 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 pe-

riod, 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.

## 2.1 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,

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

subject to the budget constraint,

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

where  $c_t$ , consumption at time  $t$ , is not indexed by individual type  $i$  since it is equal across all agents,  $\xi_t$  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,  $\mu_t$  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,  $\pi_t$  is the inflation rate,  $\Pi_t$  is the value of profits earned by owning stock in firms, and  $\tau_t$  is the value of real lump sum taxes. The preference parameters are  $\sigma \in (0, \infty)$ , which is the pseudo intertemporal-elasticity of substitution,<sup>1</sup>  $\eta \in [0, 1)$ , 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 that the first order conditions for the consumer lead to the log-linear Euler equation,

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

---

<sup>1</sup>When there is no habit formation and labor supply is fixed,  $\sigma$  is exactly equal to the intertemporal elasticity of substitution.

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

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

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

$$\hat{\xi}_t = \rho_\xi \hat{\xi}_{t-1} + \epsilon_{\xi,t}, \quad (5)$$

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

When there is no habit formation, equations (3) and (4) lead to the standard IS equation,

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

Habit formation is added to the model, because as equation (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 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. The production function is given by,

$$y_t = \left[ \int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}} \quad (6)$$

---

<sup>2</sup>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.

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

$$y_t(i) = \left[ \frac{p_t(i)}{p_t} \right]^{-\theta} y_t, \quad (7)$$

where  $p_t(i)$  is the price of intermediate good  $i$  and  $p_t$  is the price of the final good. Substituting equation (7) into (6) leads to a consumption price index that holds in equilibrium,

$$p_t = \left[ \int_0^1 p_t(i)^{1-\theta} di \right]^{\frac{1}{1-\theta}}. \quad (8)$$

### 2.2.1 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,

$$y_t(i) = z_t k_t(i)^\alpha n_t(i)^{1-\alpha} \quad (9)$$

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,

$$C_t = \frac{w_t(i)}{p_t} n_t(i) + \rho_t(i) k_t(i), \quad (10)$$

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,

$$\hat{y}_t = \hat{z}_t + \alpha \hat{k}_t + (1 - \alpha) \hat{n}_t. \quad (11)$$

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,

$$\hat{s}_t = \frac{\alpha + \mu}{1 - \alpha} \hat{y}_t - \frac{\alpha(\mu + 1)}{1 - \alpha} \hat{k}_t - \hat{\lambda}_t - \frac{\mu + 1}{1 - \alpha} \hat{z}_t, \quad (12)$$

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,

$$\hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t}, \quad (13)$$

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

### 2.2.2 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 is 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,  $\iota_t$ , 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,

$$k_{t+1}(i) = (1 - \delta)k_t(i) + \iota_t I_t(i) - \frac{\phi}{2} \left[ \frac{k_{t+1}(i)}{k_t(i)} - 1 \right]^2 k_t(i) \quad (14)$$

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 (14) then integrating across all the firms leads to the following relationship between capital and investment:

$$\hat{k}_{t+1} = (1 - \delta)\hat{k}_t + \delta\hat{I}_t + \delta\hat{\iota}_t \quad (15)$$

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

$$E_0 \sum_{t=0}^{\infty} \beta^t \lambda_t [\rho_t(i) k_t(i) - I_t(i)], \quad (16)$$

subject to equation (14). The appendix shows that profit maximization leads to the following evolution of the aggregate capital stock:

$$\begin{aligned} \hat{\lambda}_t + \phi (\hat{k}_{t+1} - \hat{k}_t) &= \beta(1 - \delta) E_t \hat{\lambda}_{t+1} + \left( \frac{1 - \beta(1 - \delta)}{1 - \alpha} \right) [(\mu + 1) E_t \hat{y}_{t+1} - (1 + \mu\alpha) \hat{k}_{t+1}] \\ &+ \beta\phi (E_t \hat{k}_{t+2} - \hat{k}_{t+1}) - \frac{(\mu + 1) [1 - \beta(1 - \delta)]}{1 - \alpha} E_t \hat{z}_{t+1} + \hat{\iota}_t - \beta(1 - \delta) E_t \hat{\iota}_{t+1} \\ &+ [1 - \beta(1 - \delta)] E_t \hat{\mu}_{t+1}, \end{aligned} \quad (17)$$

The investment shock is assumed to follow the stochastic process,

$$\hat{\iota}_t = \rho_{\iota} \hat{\iota}_{t-1} + \epsilon_{\iota,t} \quad (18)$$

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

### 2.2.3 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,  $\gamma$ , 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,  $\omega^T$  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:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left( \frac{p_{t+T}(i)}{p_{t+T}} \right) y_{t+T}(i) - S[y_{t+T}(i)] \right\}, \quad (19)$$

where  $S[y_{t+T}(i)]$  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,

$$\log p_{t+T}(i) = \log p_{t+T-1}(i) + \gamma\pi_{t+T-1} \quad (20)$$

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,

$$\pi_t = \left( \frac{1}{1 + \beta\gamma} \right) (\gamma\pi_{t-1} + \beta E_t \pi_{t+1} + \kappa \hat{s}_t) \quad (21)$$

where  $\kappa$  decreases as  $\omega$ , the degree of price stickiness, increases. The parameter  $\kappa$  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.

### 2.2.4 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,

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r) (\psi_\pi E_t \pi_{t+1} + \psi_y E_t \hat{y}_{t+1}) + \epsilon_{r,t} \quad (22)$$

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  $\sigma_r^2$ .

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.

## 2.3 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 ( $\pi_t$ ), and the interest rate ( $\hat{r}_t$ ). The demand side of the model consists of the Euler equation, (3), and the definition of the marginal utility of income, (4). The supply side of the model consists of the Phillips curve, (21), the definition of the marginal cost, (12), the evolution of capital, (17), the relationship between investment and capital, (15), and the production

function (11). The model is completed with the monetary policy rule, (22), and the following log-linear goods market clearing condition,

$$\hat{y}_t = c_y \hat{c}_t + \delta k_y \hat{I}_t, \quad (23)$$

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,

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

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

There are five exogenous shocks in the model: the preference shock,  $\xi_t$ , whose evolution is given in equation (5); the technology shock,  $z_t$ , whose evolution is given in equation (13); the investment shock,  $\iota_t$ , whose evolution is given in equation (18); the labor supply shock,  $\mu_t$ , and the monetary policy shock,  $\epsilon_{r,t}$ .

### 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 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 the 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 deter-

mine 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:

$$\Omega_0 x_t = \Omega_1 x_{t-1} + \Omega_2 E_t^* x_{t+1} + \Psi v_t, \quad (24)$$

$$v_t = A v_{t-1} + \epsilon_t \quad (25)$$

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  $\epsilon_t$  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,  $v_t = [\hat{z}_t \ \hat{l}_t \ \hat{\xi}_t \ \hat{\mu}_t \ \epsilon_{r,t}]$

The solution of the model can be written as,

$$x_t = G x_{t-1} + H v_t. \quad (26)$$

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 (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 (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 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, (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  $\Phi_t$  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 (26) by ordinary least squares, they form the estimate,

$$\Phi'_t = \left( \frac{1}{t-1} \sum_{\tau=2}^t X_\tau X'_\tau \right)^{-1} \left( \frac{1}{t-1} \sum_{\tau=2}^t X_\tau Y'_\tau \right). \quad (27)$$



The ordinary least squares estimate  $\Phi_t$  can be rewritten into the convenient recursive form:

$$\Phi_t = \Phi_{t-1} + g_t(Y_t - \Phi_{t-1}X_t)X_t'R_t^{-1}, \quad (28)$$

$$R_t = R_{t-1} + g_t(X_tX_t' - R_{t-1}), \quad (29)$$

where  $g_t = 1/(t-1)$  is the learning gain.<sup>3</sup> The recursive form demonstrates precisely how expectations are adaptive. Agents take the previous period's estimates,  $\Phi_{t-1}$  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  $\Phi$  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  $\Phi_t$  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  $\Phi_t$ , and let  $\hat{G}_t$  and  $\hat{H}_t$  denote the time  $t$  estimate of  $G$  and  $H$ , respectively, obtained from  $\Phi_t$ . Agents' expectation of  $x_{t+1}$  is given by,

$$E_t^*x_{t+1} = \hat{g}_{0,t} + \hat{G}_tE_t^*x_t + \hat{H}_tE_tv_{t+1} \quad (30)$$

Note that equation (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

---

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

agents, agents must also estimate  $x_t$  by least squares. The time  $t$  estimate of  $x_t$  is given by,

$$E_t x_t^* = \hat{g}_{0,t} + \hat{G}_t x_{t-1} + \hat{H}_t v_t. \quad (31)$$

Plugging this into equation (30) yields,

$$E_t^* x_{t+1} = (I + \hat{G}_t) \hat{g}_{0,t} + \hat{G}_t^2 x_{t-1} + (\hat{G}_t \hat{H}_t + \hat{H}_t A) v_t. \quad (32)$$

Plugging the agents' forecast, (32), into the structural form of the model, (24), leads to the following actual law of motion for  $x_t$ :

$$x_t = \Omega_0^{-1} \Omega_2 (I + \hat{G}_t) \hat{g}_{0,t} + \Omega_0^{-1} (\Omega_1 + \Omega_2 \hat{G}_t^2) x_{t-1} + \Omega_0^{-1} [\Psi + \Omega_2 (\hat{G}_t \hat{H}_t + \hat{H}_t A)] v_t. \quad (33)$$

## 4 Estimation

### 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,

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

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.

## 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  $\Phi$  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 (28) and (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,

$$\Phi_t = \left( \sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} X'_{t-\tau} \right)^{-1} \left( \sum_{\tau=0}^{t-1} (1-g)^t X_{t-\tau} Y'_{t-\tau} \right) \quad (34)$$

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,

$$\hat{c}_t = \frac{CONS_t^* - c_y y^*}{c_y y^*}, \quad (35)$$

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  $\pi^*$  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  $\pi_t$  according to,

$$\pi_t = \frac{1}{400} (INF_t - \pi^*). \quad (36)$$

The steady state inflation rate,  $\pi^*$ , 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  $\beta^{-1}$ . 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

$$\hat{r}_t = \frac{1}{400} (FF_t - r^* - \pi^*). \quad (37)$$

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 (15) describes the evolution of the capital stock in terms of the percentage deviation from the steady state:

$$\hat{k}_{t+1} = (1 - \delta)\hat{k}_t + \delta\hat{I}_t + \delta\hat{l}_t. \quad (38)$$

For a given initial value for  $\hat{k}_t$  in the pre-sample, a simulated path for the investment shock,  $\hat{l}_t$ , 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,

$$\hat{I}_t = \frac{INV_t^* - (1 - c_y)y^*}{(1 - c_y)y^*}, \quad (39)$$

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 (18) given the variance of the investment shock,  $\sigma_l^2$ , 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  $\hat{c}_t$ ,  $\hat{k}_t$ ,  $\pi_t$ , and  $\hat{r}_t$  according to equations (35), (36), (37), (38), (18),

and (39). The initial condition for  $\Phi_0$  is then computed using equation for the weighted least squares procedure, equation (34).

### 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 (33), and the evolution of the structural shocks given in equation (25). Equation (33) can be rewritten more compactly as,

$$x_t = b_t + F_t x_{t-1} + M_t v_t, \quad (40)$$

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

$$\begin{aligned} b_t &= \Omega_0^{-1} \Omega_2 \left( I + \hat{G}_t \right) \hat{g}_{0,t}, \\ F_t &= \Omega_0^{-1} \left( \Omega_1 + \Omega_2 \hat{G}_t^2 \right), \\ M_t &= \Omega_0^{-1} \left[ \Psi + \Omega_2 \left( \hat{G}_t \hat{H}_t + \hat{H}_t A \right) \right] \end{aligned}$$

This equation can be combined with equation (25) into the single state equation,

$$x_t^* = b_t^* + F_t^* x_{t-1}^* + \epsilon_t^*, \quad (41)$$

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

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

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

The variance of  $\epsilon_t^*$  is given by,

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

where  $\Sigma$  is a diagonal matrix with the variance of the structural shocks along the diagonal.

The observations equations are given by,

$$\begin{aligned} CONS_t^* &= c_y y^* + c_y y^* \hat{c}_t \\ INV_t^* &= (1 - c_y) y^* + (1 - c_y) y^* \hat{I}_t \\ INF_t &= \pi^* + 400 \pi_t \\ FF_t &= \pi^* + 400 (r^n + \hat{r}_t), \end{aligned} \tag{42}$$

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,  $\beta$ , is set equal to 0.9925 which corresponds to a steady state real interest rate approximately equal to 3%. The depreciation rate,  $\delta$ , is set equal to 0.025 which corresponds to an approximate annual depreciation rate of 10%.

The steady state level of inflation,  $\pi^*$ , 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  $\kappa$ , the slope on the

Phillips curve that depends on the degree of price flexibility, and  $\alpha$ , 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  $\rho_\mu$  and  $\sigma_\mu$  out of the estimation.

## 5 Results

### 5.1 Parameter Estimates

The parameter estimates for the four cases for expectations formation are given in Table 1. In this subsection I look at each case in turn.

#### Case 1: Rational Expectations

The first two columns of Table 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 (17). This equation illustrates that large values for  $\mu$  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 (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 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 ( $\sigma$ ), the inverse elasticity of labor supply ( $\mu$ ), and the cost of capital adjustment ( $\phi$ ), 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 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  $\mu$ , but possibly contradicts the finding that the estimate for  $\sigma$  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.

## 5.2 Performance Comparison

The top part of Table 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 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 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 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 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,  $\Phi_t$  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.

### 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 3, 4, 5, and 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  $\Phi_t$  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 3, 4, 5, and 6 show the difference between the evolution of the variables after the shock and the evolution the variables would take in the absence of any shocks.

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 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 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 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 7 shows the smoothed estimates<sup>4</sup> 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 its 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 4 that in Case 2, technology shocks cause long-lived oscillatory behavior in consumption, investment, and inflation, whereas under rational expectations and Case 3, the responses of these variables are in one direction and long lived. Despite the

---

<sup>4</sup>Kalman filter smoothing is computed using the techniques described in de Jong (1989).

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 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 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 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.

## 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 Policymaker 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 York 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 Dynamics,” 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 Association, 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.

## A Derivation of New Keynesian model

### 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 (1), subject to the budget constraint, given in equation (2). The first order conditions are,

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

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

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

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

$$\hat{\lambda}_t = \frac{1}{\sigma(1 - \beta\eta)(1 - \eta)} \left[ \beta\eta E_t \hat{c}_{t+1} - (1 + \beta\eta^2) \hat{c}_t + \eta \hat{c}_{t-1} \right] + \left( \hat{\xi}_t - \beta\eta E_t \hat{\xi}_{t+1} \right) \quad (43)$$

$$\hat{w}_t(i) - \hat{p}_t = \mu \hat{n}_t(i) - \hat{\lambda}_t + \hat{\mu}_t \quad (44)$$

$$\hat{\lambda}_t = E_t \hat{\lambda}_{t+1} + \hat{r}_t - E_t \pi_{t+1} \quad (45)$$

where a hat indicates the percentage deviation of the variable from its steady state. Equation (44) will be referenced later to express equilibrium real wages in terms of employment. Equations (45) and (43) 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  $i$ ,

$$y_t(i) = \left[ \frac{p_t(i)}{p_t} \right]^{-\theta} y_t. \quad (46)$$

which is given in equation (7).

### A.2.2 Input choices

Intermediate goods firms choose labor demand and rent capital to minimize real total cost, given in equation (10), subject to the production function, given in equation (9). The first order conditions are,

$$\frac{w_t(i)}{p_t} = (1 - \alpha) s_t(i) \frac{y_t(i)}{n_t(i)}, \quad (47)$$

$$\rho_t(i) = \alpha s_t(i) \frac{y_t(i)}{k_t(i)}, \quad (48)$$

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,

$$\hat{\rho}_t(i) = \hat{s}_t(i) + \hat{y}_t(i) - \hat{k}_t(i), \quad (49)$$

$$\hat{w}_t(i) - \hat{p}_t = \hat{s}_t(i) + \hat{y}_t(i) - \hat{n}_t(i), \quad (50)$$



Combining these two equations to eliminate  $\hat{s}_t(i)$  and substituting equation (44) to eliminate wages and prices leads to the expression for the rental rate of capital,

$$\hat{\rho}_t(i) = (\mu + 1) \hat{n}_t(i) - \hat{k}_t(i) - \hat{\lambda}_t + \hat{\mu}_t. \quad (51)$$

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,

$$\hat{y}_t(i) = \hat{z}_t + \alpha \hat{k}_t(i) + (1 - \alpha) \hat{n}_t(i). \quad (52)$$

Solving equation (52) for  $\hat{n}_t(i)$  and substituting this into (51) yields,

$$\hat{\rho}_t(i) = \frac{\mu + 1}{1 - \alpha} \hat{y}_t(i) - \frac{1 + \mu\alpha}{1 - \alpha} \hat{k}_t(i) - \hat{\lambda}_t + \hat{\mu}_t - \frac{\mu + 1}{1 - \alpha} \hat{z}_t \quad (53)$$

Solving equation (49) for  $\hat{s}_t(i)$  and using equation (53) to substitute out  $\hat{\rho}_t(i)$  leads to the expression for marginal cost for firm  $i$ ,

$$\hat{s}_t(i) = \frac{\alpha + \mu}{1 - \alpha} \hat{y}_t(i) - \frac{1 + \alpha\mu}{1 - \alpha} \hat{k}_t(i) - \hat{\lambda}_t + \hat{\mu}_t - \frac{\mu + 1}{1 - \alpha} \hat{z}_t \quad (54)$$

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

$$\hat{s}_t = \frac{\alpha + \mu}{1 - \alpha} \hat{y}_t - \frac{\alpha(\mu + 1)}{1 - \alpha} \hat{k}_t - \hat{\lambda}_t + \hat{\mu}_t - \frac{\mu + 1}{1 - \alpha} \hat{z}_t \quad (55)$$

Subtracting equation (55) from equation (54), 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,

$$\hat{s}_t(i) = \hat{s}_t + \frac{\alpha + \mu}{1 - \alpha} [\hat{y}_t(i) - \hat{y}_t] - \frac{\alpha + \mu}{1 - \alpha} \tilde{k}_t(i) \quad (56)$$

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 (16), subject to the evolution of firm-specific capital stock, given in equation (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,

$$\begin{aligned} \frac{\lambda_t}{\iota_t} \left[ 1 + \phi \left( \frac{k_{t+1}(i)}{k_t(i)} - 1 \right) \right] = \\ \beta E_t \frac{\lambda_{t+1}}{\iota_{t+1}} \left[ \iota_{t+1} \rho_{t+1}(i) + (1 - \delta) + \phi \left( \frac{k_{t+2}(i)}{k_{t+1}(i)} - 1 \right) \frac{k_{t+2}(i)}{k_{t+1}(i)} - \frac{\phi}{2} \left( \frac{k_{t+2}(i)}{k_{t+1}(i)} - 1 \right)^2 \right]. \end{aligned} \quad (57)$$

Log-linearizing this yields,

$$\begin{aligned} \hat{\lambda}_t + \phi \left( \hat{k}_{t+1}(i) - \hat{k}_t(i) \right) &= E_t \hat{\lambda}_{t+1} + [1 - \beta(1 - \delta)] E_t \hat{\rho}_{t+1}(i) \\ &+ \beta \phi \left( E_t \hat{k}_{t+2}(i) - \hat{k}_{t+1}(i) \right) + \hat{\iota}_t - \beta(1 - \delta) E_t \hat{\iota}_{t+1}. \end{aligned} \quad (58)$$

Plugging equation (53) into (58) leads to the following equilibrium condition for the evolution of capital stock for firm  $i$ :

$$\begin{aligned} \hat{\lambda}_t + \phi \left( \hat{k}_{t+1}(i) - \hat{k}_t(i) \right) &= \beta(1 - \delta) E_t \hat{\lambda}_{t+1} \\ &+ \left( \frac{1 - \beta(1 - \delta)}{1 - \alpha} \right) \left[ (\mu + 1) E_t \hat{y}_{t+1}(i) - (1 + \mu\alpha) \hat{k}_{t+1}(i) \right] + \beta \phi \left( E_t \hat{k}_{t+2}(i) - \hat{k}_{t+1}(i) \right) \\ &+ [1 - \beta(1 - \delta)] E_t \hat{\mu}_{t+1} - \frac{(\mu + 1) [1 - \beta(1 - \delta)]}{1 - \alpha} E_t \hat{z}_{t+1} + \hat{\iota}_t - \beta(1 - \delta) E_t \hat{\iota}_{t+1}. \end{aligned} \quad (59)$$

Integrating equation (59) over all firms leads to the evolution of the aggregate capital stock,

$$\begin{aligned} \hat{\lambda}_t + \phi(\hat{k}_{t+1} - \hat{k}_t) &= \beta(1 - \delta)E_t\hat{\lambda}_{t+1} + \left(\frac{1 - \beta(1 - \delta)}{1 - \alpha}\right) [(\mu + 1)E_t\hat{y}_{t+1} - (1 + \mu\alpha)\hat{k}_{t+1}] \\ &+ \beta\phi(E_t\hat{k}_{t+2} - \hat{k}_{t+1}) - \frac{(\mu + 1)[1 - \beta(1 - \delta)]}{1 - \alpha}E_t\hat{z}_{t+1} + \hat{\iota}_t - \beta(1 - \delta)E_t\hat{\iota}_{t+1} \\ &+ [1 - \beta(1 - \delta)]E_t\hat{\mu}_{t+1}, \end{aligned} \quad (60)$$

which is shown in equation (17) of the paper. Subtracting equation (60) from equation (59) leads to following expression for firm  $i$ 's capital stock in terms of the aggregate capital stock,

$$\begin{aligned} \phi(\tilde{k}_{t+1}(i) - \tilde{k}_t(i)) &= \beta\phi(E_t\tilde{k}_{t+2}(i) - \tilde{k}_{t+1}(i)) \\ &+ \left[\frac{1 - \beta(1 - \delta)}{1 - \alpha}\right] [(\mu + 1)E_t(\hat{y}_{t+1}(i) - \hat{y}_{t+1}) - (1 + \mu\alpha)\tilde{k}_{t+1}(i)]. \end{aligned} \quad (61)$$

#### A.2.4 Optimal pricing

The inflation indexation rule given in equation (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 (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, (7), into the profit function, (19), to express the profit only in terms of the intermediate good price,  $p_t(i)$ , and aggregate state variables the firm cannot control:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ \left( \frac{p_t(i)e^{\gamma\pi_{t+T}^*}}{p_{t+T}} \right)^{1-\theta} y_{t+T} - S \left[ \left( \frac{p_t(i)e^{\gamma\pi_{t+T}^*}}{p_{t+T}} \right)^{-\theta} y_{t+T} \right] \right\}. \quad (62)$$

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

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \frac{\lambda_{t+T}}{\lambda_t} \left\{ (1-\theta) \left( \frac{p_t^*(i)e^{\gamma\pi_{t+T}^*}}{p_{t+T}} \right)^{1-\theta} + \theta s_{t+T}(i) \left( \frac{p_t^*(i)e^{\gamma\pi_{t+T}^*}}{p_{t+T}} \right)^{-\theta} \right\} \frac{y_{t+T}}{p_t^*(i)} = 0, \quad (63)$$

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 (63) as,

$$\begin{aligned} (1-\theta) E_t \sum_{T=0}^{\infty} (\omega\beta)^T \lambda_{t+T} \left( \frac{p_t^*(i)e^{\gamma\pi_{t+T}^*}}{p_{t+T}} \right)^{1-\theta} y_{t+T} = \\ -\theta E_t \sum_{T=0}^{\infty} (\omega\beta)^T \lambda_{t+T} s_{t+T}(i) \left( \frac{p_t^*(i)e^{\gamma\pi_{t+T}^*}}{p_{t+T}} \right)^{-\theta} y_{t+T}, \end{aligned} \quad (64)$$

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

$$(1-\theta) \lambda y E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[ \hat{\lambda}_{t+T} + \hat{y}_{t+T} + (1-\theta) \left( \hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* \right) \right], \quad (65)$$

$$-\theta \lambda y s E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[ \hat{\lambda}_{t+T} + \hat{y}_{t+T} + \hat{s}_{t+T} - \theta \left( \hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* \right) \right] \quad (66)$$

where  $\lambda$  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 (63) where  $\lambda_t = \lambda$  and  $p_t^*(i) = p_t = p$  for all  $t$ . In the steady state equation (63) 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,

$$s = -\frac{1 - \theta}{\theta}. \quad (67)$$

The coefficient  $-\theta s$  in equation (66) therefore cancels out with  $1 - \theta$  in equation (65). Combining the left and right hand side then yields,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[ \hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* - \hat{s}_{t+T}(i) \right] = 0 \quad (68)$$

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

$$\hat{p}_t^*(i) = (1 - \omega\beta) E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[ \hat{p}_{t+T} - \gamma\pi_{t+T}^* + \hat{s}_{t+T}(i) \right]. \quad (69)$$

Substitute into equation (69), the log-linearized the demand for intermediate good  $i$  at time  $t + T$ , which is given by,

$$\hat{y}_{t+T}(i) = -\theta(\hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^*) + \hat{y}_{t+T} \quad (70)$$

and the marginal cost given in equation (56). This leads to an expression for the optimal price for firm  $i$  in terms of aggregate variables and the firm's expected future capital,

$$\begin{aligned} \hat{p}_t^*(i) = & (1 - \omega\beta) E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left\{ \hat{p}_{t+T} - \gamma\pi_{t+T}^* + \hat{s}_{t+T} - \frac{\theta(\alpha + \mu)}{1 - \alpha} \left[ \hat{p}_t^*(i) - \hat{p}_{t+T} + \gamma\pi_{t+T}^* \right] \right\} \\ & - (1 - \omega\beta) E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left\{ \frac{\alpha(\mu + 1)}{1 - \alpha} \tilde{k}_{t+T}(i) \right\}. \end{aligned} \quad (71)$$

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

$$\hat{p}_t^*(i) = (1 - \omega\beta) E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left[ \hat{p}_{t+T} - \gamma\pi_{t+T}^* + \psi\hat{s}_{t+T} - \frac{\psi\alpha(\mu + 1)}{1 - \alpha} \tilde{k}_{t+T}(i) \right], \quad (72)$$

where

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

Equation (72) can be rewritten as the first order difference equation:

$$\hat{p}_t^*(i) = \omega\beta E_t \hat{p}_{t+1}^*(i) + (1 - \omega\beta) \left( \hat{p}_t + \psi \hat{s}_t - \frac{\psi\alpha(\mu + 1)}{\mu(1 - \alpha)} \tilde{k}_t(i) \right), \quad (73)$$

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  $\omega$  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,

$$E_t \hat{p}_{t+1}(i) = \omega [\hat{p}_t^*(i) + \gamma\pi_{t-1}] + (1 - \omega) E_t \hat{p}_{t+1}^*(i). \quad (74)$$

### 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 (73) 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 (71) 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:

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) \quad (75)$$

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 (61). Substituting the log-linear demand for  $y_{t+1}(i)$  into equation (61) leads to,

$$\begin{aligned} \phi\left(\tilde{k}_{t+1}(i) - \tilde{k}_t(i)\right) &= \beta\phi\left(E_t\tilde{k}_{t+2}(i) - \tilde{k}_{t+1}(i)\right) \\ &- \left[\frac{1 - \beta(1 - \delta)}{1 - \alpha}\right] \left[\theta(\mu + 1)E_t\tilde{p}_{t+1}(i) - (1 + \mu\alpha)\tilde{k}_{t+1}(i)\right]. \end{aligned} \quad (76)$$

where  $\tilde{p}_{t+1}(i) = \hat{p}_{t+1}(i) - \hat{p}_t(i)$  is the relative price of intermediate good  $i$  in period  $t + 1$ . The rational expectations solution for (76) must have the form,

$$\tilde{k}_{t+1}(i) = m\tilde{k}_t(i) + n\tilde{p}_t(i), \quad (77)$$

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

$$\tilde{k}_{t+1}(i) = m\tilde{k}_t(i) + n\hat{p}_t^*(i) - n\hat{p}_t(i). \quad (78)$$

Substituting this into equation (75) shows that,

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

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,

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

Solving this equation yields,

$$E_t \sum_{T=0}^{\infty} (\omega\beta)^T \tilde{k}_{t+T}(i) = \frac{1}{1 - \omega\beta m} \left[ \frac{\omega\beta n}{1 - \omega\beta} \hat{p}_t^*(i) + \tilde{k}_t(i) - \omega\beta n E_t \sum_{T=0}^{\infty} (\omega\beta)^T \hat{p}_{t+T}(i) \right]. \quad (79)$$

Substituting this into equation (71) and solving for  $\hat{p}_t^*(i)$  leads to the following solution,

$$\hat{p}_t^*(i) = (1 - \omega\beta)E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left( \hat{p}_{t+T} - \gamma\pi_{t+T}^* + \nu\hat{s}_{t+T} \right) - \frac{\alpha\nu(\mu+1)(1-\omega\beta)}{(1-\alpha)(1-\omega\beta m)} \tilde{k}_t(i), \quad (80)$$

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 (80) 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 (80) implies,

$$\hat{p}_t^* = (1 - \omega\beta)E_t \sum_{T=0}^{\infty} (\omega\beta)^T \left( \hat{p}_{t+T} - \gamma\pi_{t+T}^* + \nu\hat{s}_{t+T} \right). \quad (81)$$

This can be rewritten as the first order difference equation,

$$\hat{p}_t^* = \omega\beta E_t \hat{p}_{t+1}^* + (1 - \omega\beta) (\hat{p}_t + \nu\hat{s}_t). \quad (82)$$

Substituting equation (81) into (82) to eliminate  $\hat{p}_t^*$  and  $E_t \hat{p}_{t+1}^*$  leads to the Phillips curve,

$$\pi_t = \left( \frac{1}{1 + \beta\gamma} \right) [\gamma\pi_{t-1} + \beta E_t \pi_{t+1} + \kappa\hat{s}_t], \quad (83)$$

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 (78) which must satisfy the optimality condition for capital given in equation (76). Equation (76) can be rearranged as,

$$\tilde{k}_{t+1}(i) = \tilde{k}_t(i) + \beta E_t \tilde{k}_{t+2}(i) - \zeta_0 E_t \tilde{p}_{t+1}(i) - \zeta_1 \tilde{k}_{t+1}(i), \quad (84)$$

where  $\zeta_0$  and  $\zeta_1$  are given by,

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

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

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 (74), the expected relative price can be rewritten as,

$$E_t \tilde{p}_{t+1}(i) = E_t \hat{p}_{t+1}(i) - E_t \hat{p}_{t+1} = \omega \hat{p}_t(i) + (1 - \omega) E_t \hat{p}_{t+1}^*(i) - E_t \hat{p}_{t+1} \quad (85)$$

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 (73), the rational expectation solution for  $\hat{p}_t^*(i)$  must take the form,

$$\hat{p}_t^*(i) = f(\hat{p}_t, \hat{s}_t) + a \tilde{k}_t(i), \quad (86)$$

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 (86) can be rewritten as,

$$\hat{p}_t^*(i) = \hat{p}_t^* + a\tilde{k}_t(i). \quad (87)$$

Advancing equation (87) one period and taking expectations yields,

$$E_t \hat{p}_{t+1}^*(i) = E_t \hat{p}_{t+1}^* + a\tilde{k}_{t+1}(i), \quad (88)$$

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  $\omega$  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,

$$E_t \hat{p}_{t+1} = \omega \hat{p}_t + (1 - \omega) E_t \hat{p}_{t+1}^*. \quad (89)$$

Solving (89) for  $E_t \hat{p}_{t+1}^*$  and substituting this expression into (88) leads to,

$$E_t \hat{p}_{t+1}^*(i) = \frac{1}{1 - \omega} (E_t \hat{p}_{t+1} - \omega \hat{p}_t) + a\tilde{k}_{t+1}(i). \quad (90)$$

Substituting equation (78) for  $\tilde{k}_{t+1}(i)$  yields,

$$E_t \hat{p}_{t+1}^*(i) = \frac{1}{1 - \omega} (E_t \hat{p}_{t+1} - \omega \hat{p}_t) + am\tilde{k}_t(i) - an\tilde{p}_t(i). \quad (91)$$

Plugging this into equation (85) leads to an expression for  $E_t \tilde{p}_{t+1}(i)$  in terms of  $\tilde{p}_t(i)$  and

$\tilde{k}_t(i)$ ,

$$E_t \tilde{p}_{t+1}(i) = [\omega + (1 - \omega) an] \tilde{p}_t(i) + (1 - \omega) am \tilde{k}_t(i) \quad (92)$$

Next, using equation (78), the expected future capital stock is given by,

$$E_t \tilde{k}_{t+2}(i) = m^2 \tilde{k}_t(i) + mn \tilde{p}_t(i) + n E_t \tilde{p}_{t+1}(i) \quad (93)$$

Substituting equation (92) into equation (93) leads to an expression for  $E_t \tilde{k}_{t+2}(i)$  in terms of  $\tilde{p}_t(i)$  and  $\tilde{k}_t(i)$ ,

$$E_t \tilde{k}_{t+2}(i) = [m^2 + amn(1 - \omega)] \tilde{k}_t(i) + [mn + n\omega + an^2(1 - \omega)] \tilde{p}_t(i) \quad (94)$$

Plugging in equations (92), (93), and (78) into (84) leads to an expression for capital of the form given in equation (78) where  $m$  and  $n$  must satisfy, respectively,

$$\beta m^2 + [\beta an(1 - \omega) - \zeta_1 - \zeta_0 a(1 - \omega) - 1] m + 1 = 0, \quad (95)$$

$$\beta a(1 - \omega) n^2 + [\beta m + \beta \omega - \zeta_1 - \zeta_0 a(1 - \omega) - 1] n - \zeta_0 \omega = 0. \quad (96)$$

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 (91) into equation (73) and solving for  $\hat{p}_t^*(i)$  yields,

$$\begin{aligned} \hat{p}_t^*(i) = & \frac{1}{1 - \omega \beta an} \left( \omega \beta am - \frac{\psi \alpha (\mu + 1)}{(1 - \alpha)} \right) \tilde{k}_t(i) \\ & + \frac{\omega \beta}{(1 - \omega)(1 - \omega \beta an)} (E_t \hat{p}_{t+1} - \omega \hat{p}_t) + \hat{p}_t + \frac{\psi}{1 - \omega \beta an} \hat{s}_t, \end{aligned} \quad (97)$$

which implies  $a$  must satisfy the quadratic equation,

$$\omega \beta na^2 + (\omega \beta m - 1) a - \frac{\alpha \psi (\mu + 1)}{1 - \alpha} = 0. \quad (98)$$

Equations (95), (96), and (98) 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 (77) and (92) can be rewritten as the following dynamic system:

$$\begin{bmatrix} \tilde{k}_{t+1}(i) \\ E_t \tilde{p}_{t+1}(i) \end{bmatrix} = \begin{bmatrix} m & n \\ \omega + (1 - \omega)an & 1 - \omega \end{bmatrix} \begin{bmatrix} \tilde{k}_t(i) \\ \tilde{p}_t(i) \end{bmatrix}. \quad (99)$$

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 (99) are inside the unit circle. The eigenvalues are given by,

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

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

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,

$$\sqrt{[m + \omega + (1 - \omega)an]^2 - 4m\omega} < 2 - m - \omega - (1 - \omega)an. \quad (100)$$

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,

$$[m + \omega + (1 - \omega)an]^2 - 4m\omega < 4 - 4[m + \omega + (1 - \omega)an] + 4[m + \omega + (1 - \omega)an]^2 \quad (101)$$

This inequality does not preserve the restriction implied in (100) that the right hand side be

positive. Therefore (100) also implies

$$2 - m - \omega - (1 - \omega)an > 0. \quad (102)$$

The inequalities (101) and (102) simplify to, respectively,

$$m < 1 - an \quad (103)$$

$$m < 1 + (1 - \omega)(1 - an) \quad (104)$$

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,

$$m > -1 - \frac{1 - \omega}{1 + \omega}an \quad (105)$$

Finally, the coefficients  $m$ ,  $n$ , and  $a$  can be found by the solving the system of quadratic equations (95), (96), and (98), subject to the inequalities (103), (104), and (105).

### 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,

$$\hat{y}_t = c_y \hat{c}_t + \delta k_y \hat{I}_t, \quad (106)$$

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 (48), and the steady state

first order condition for investment, given in equation (57). Evaluating equation (48) at the steady state and using the steady state marginal cost, given in equation (67), yields,

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

Evaluating equation (57) at the steady state yields,

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

Combining these equations to eliminate  $\rho$  leads to the following capital to output ratio,

$$k_y = \frac{\beta \alpha (\theta - 1)}{\theta (1 - \beta + \beta \delta)} \quad (107)$$

Evaluating the goods market clearing condition, (106), at the steady state yields the following consumption to output ratio,

$$c_y = 1 - \delta k_y. \quad (108)$$

Table 1: Maximum Likelihood Parameter Estimates: Sample 1970:Q1 - 2008:Q1

Description	Parameter	Case 1		Case 2		Case 3		Case 4	
		Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error	Estimate	Standard Error
Habit Persistence	$\eta$	0.1060	0.0272	0.1289	0.0399	0.1224	0.0264	0.2728	0.0232
Intertemporal Substitution	$\sigma$	0.1603	0.0342	0.0513	0.0072	0.0157	0.0001	0.1220	0.0175
Inverse Elas. Labor Supply	$\mu$	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	$\phi$	7.6832	1.4003	24.8826	1.4036	26.8332	4.1201	26.9755	1.7912
Price Indexation	$\gamma$	0.3624	0.1929	0.0000	0.0349	0.5236	0.0852	0.6090	0.1214
MP Persistence	$\rho_r$	0.1945	0.0497	0.6592	0.0141	0.7250	0.0280	0.0956	0.1061
MP Output	$\psi_y$	0.0000	0.0172	0.0576	0.0090	0.0458	0.0092	0.0041	0.0131
MP Inflation	$\psi_\pi$	2.1212	0.1893	1.4448	0.0658	1.7735	0.1278	1.8491	0.0974
Preference Shock Pers.	$\rho_\xi$	0.9826	0.0069	0.9925	0.0040	0.9636	0.0070	1.0000	0.0000
Technology Shock Pers.	$\rho_z$	0.9668	0.0058	0.6741	0.0254	0.9638	0.0193	0.9506	0.0121
Investment Shock Pers.	$\rho_i$	0.9060	0.0151	0.9301	0.0084	0.9297	0.0103	0.9234	0.0173
Pref. Shock Std. Dev.	$\sigma_\xi$	0.0926	0.0252	0.1647	0.0211	0.6165	0.0511	0.6879	0.0446
Tech. Shock Std. Dev.	$\sigma_z$	0.0104	0.0003	0.0199	0.0018	0.0294	0.0037	0.0730	0.0191
Inv. Shock Std. Dev.	$\sigma_i$	0.0206	0.0021	0.1130	0.0149	0.0387	0.0016	0.0683	0.0069
Policy Shock Std. Dev.	$\sigma_r$	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

Figure 1: Forecast Errors

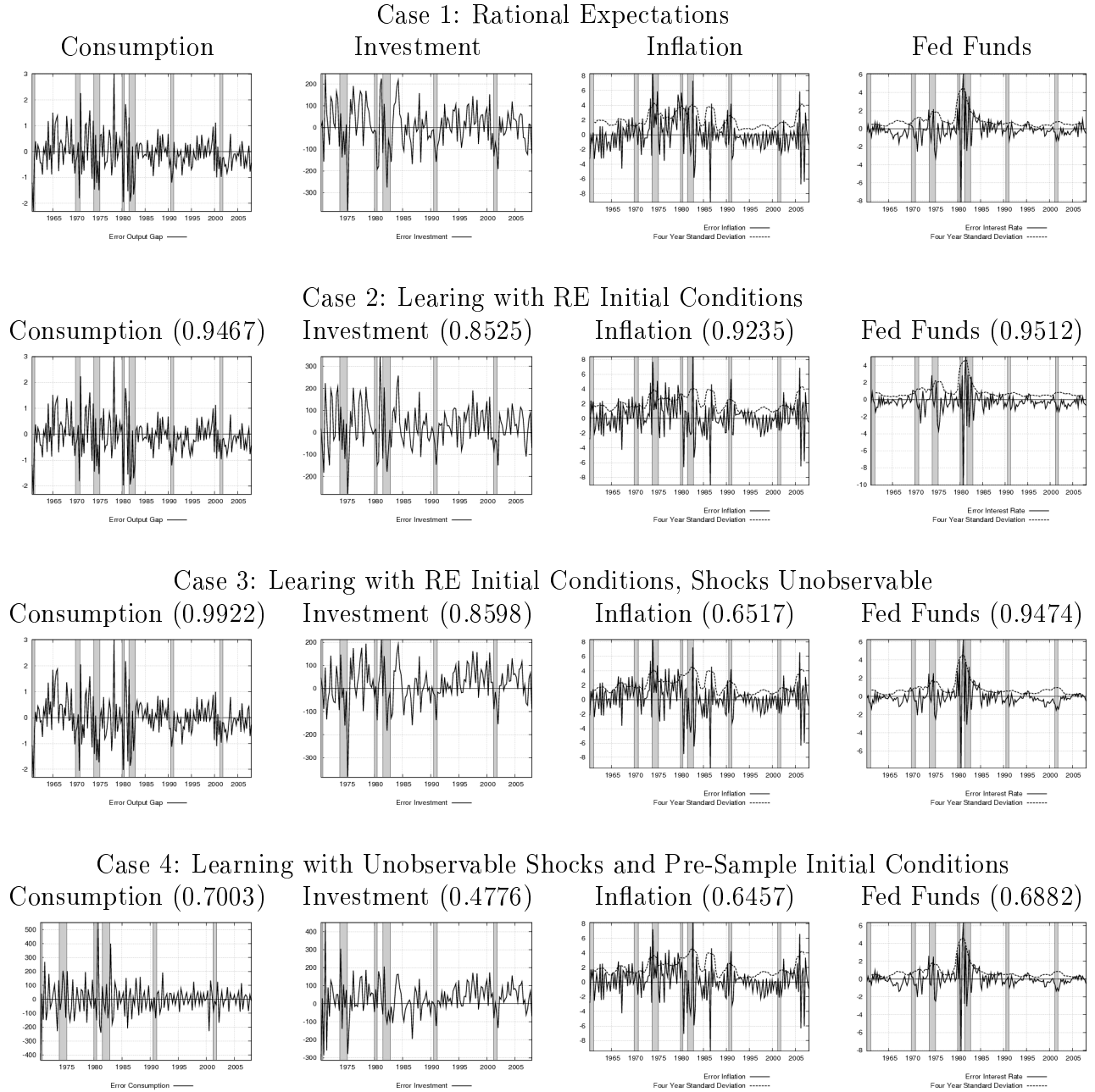




Figure 2: Out of Sample Multiperiod Forecast Errors

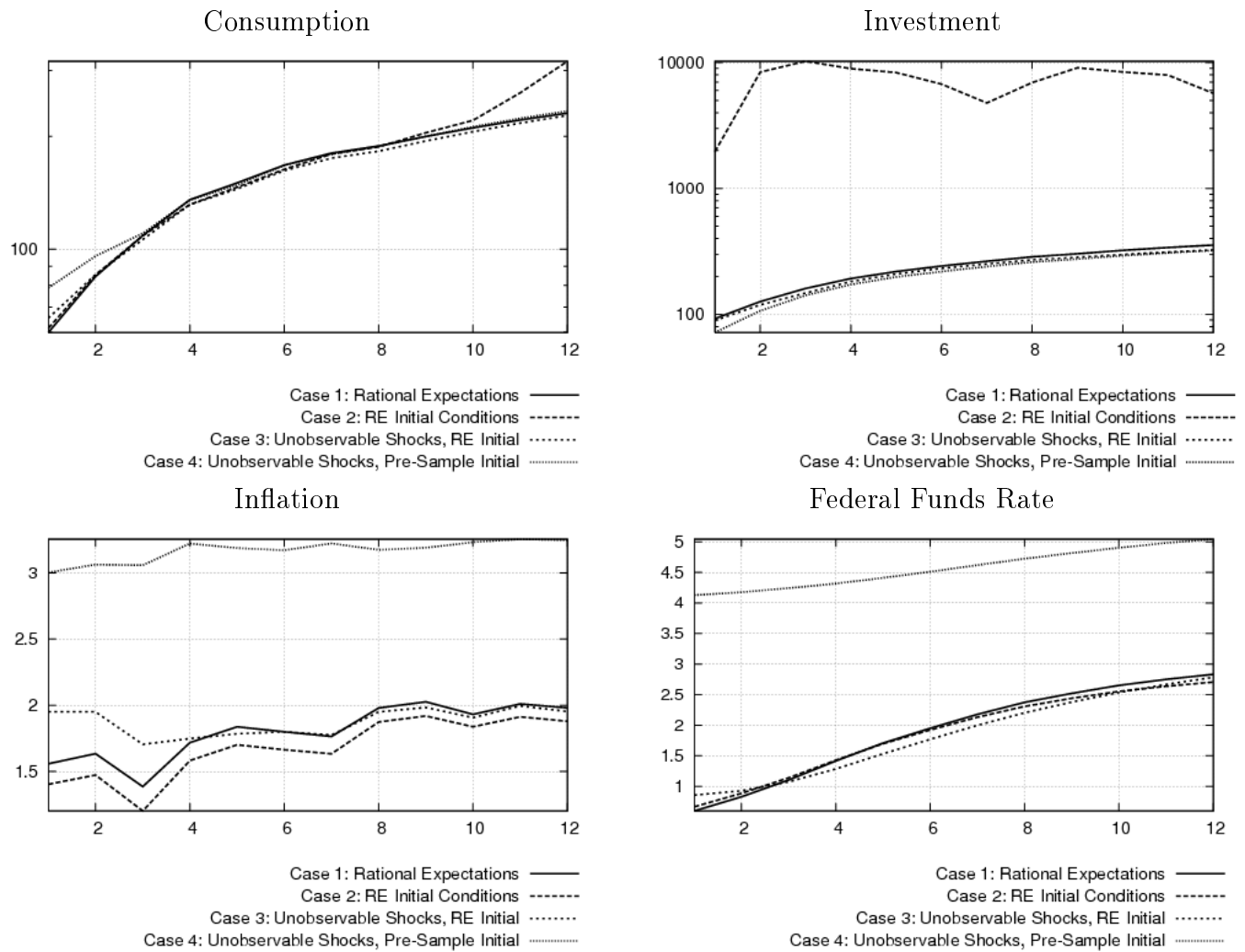


Figure 3: Preference Shock Impulse Responses

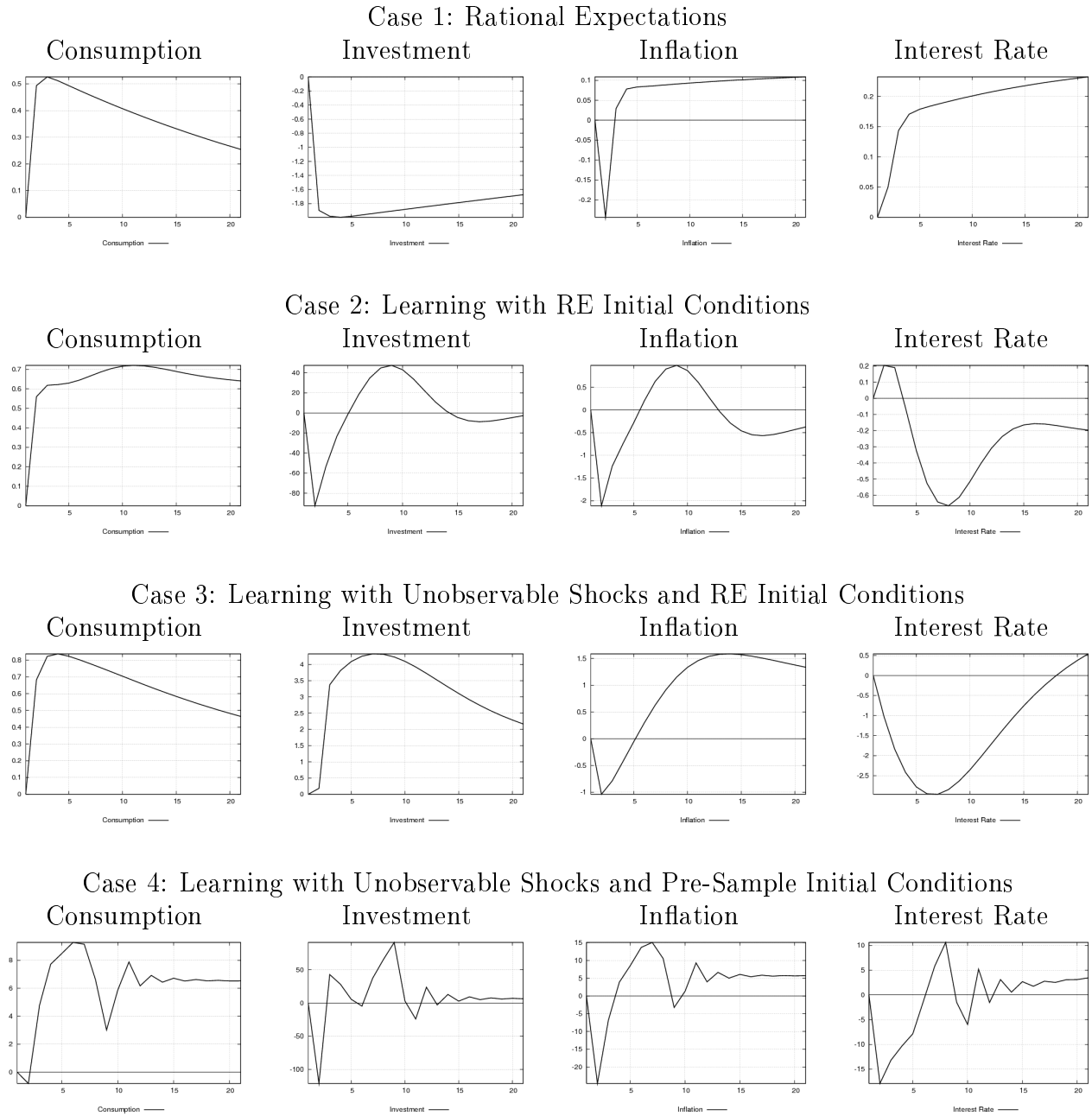


Figure 4: Technology Shock Impulse Responses

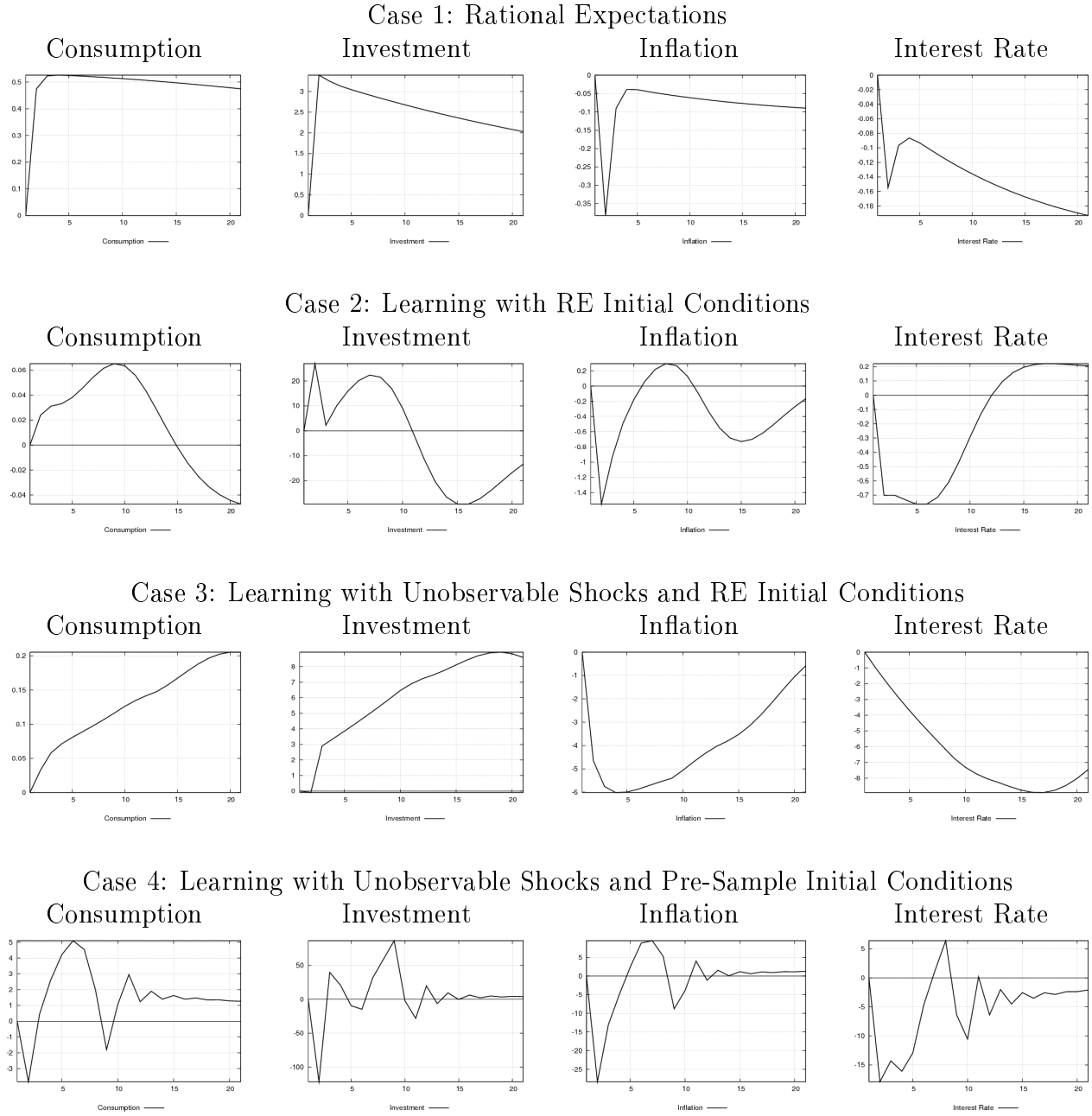


Figure 5: Investment Shock Impulse Responses

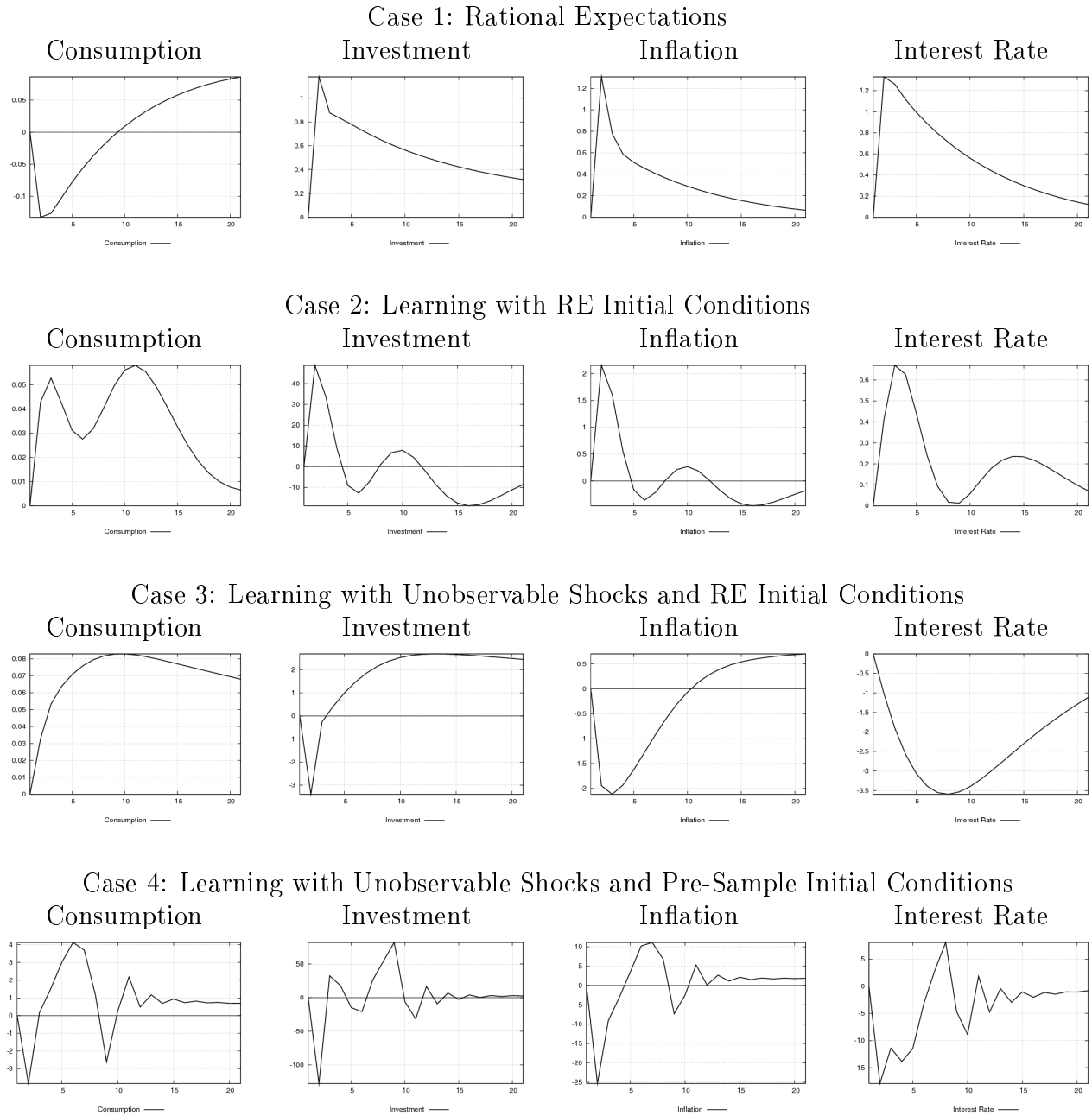


Figure 6: Monetary Policy Shock Impulse Responses

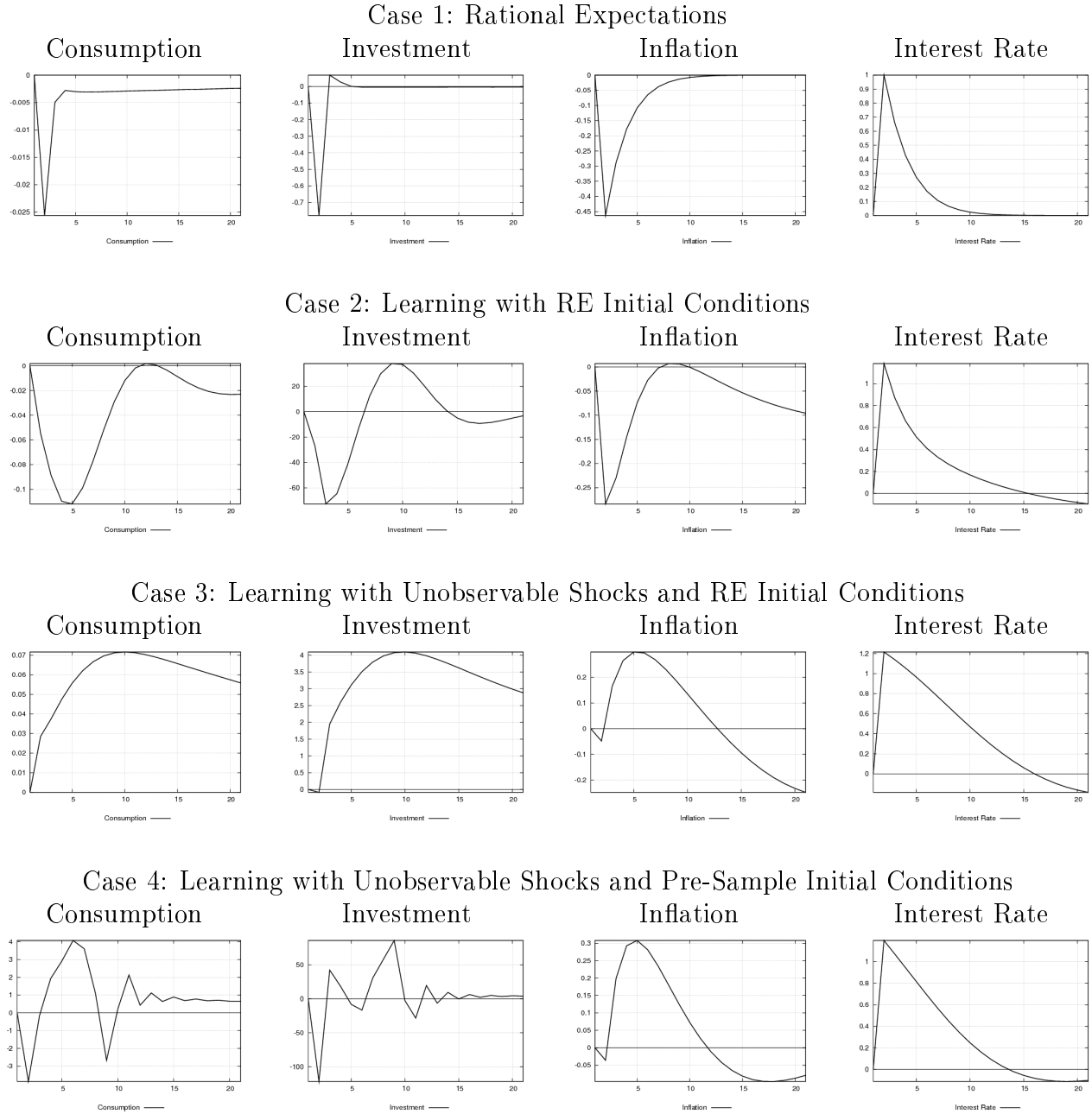


Figure 7: Smoothed Estimates of Structural Shocks

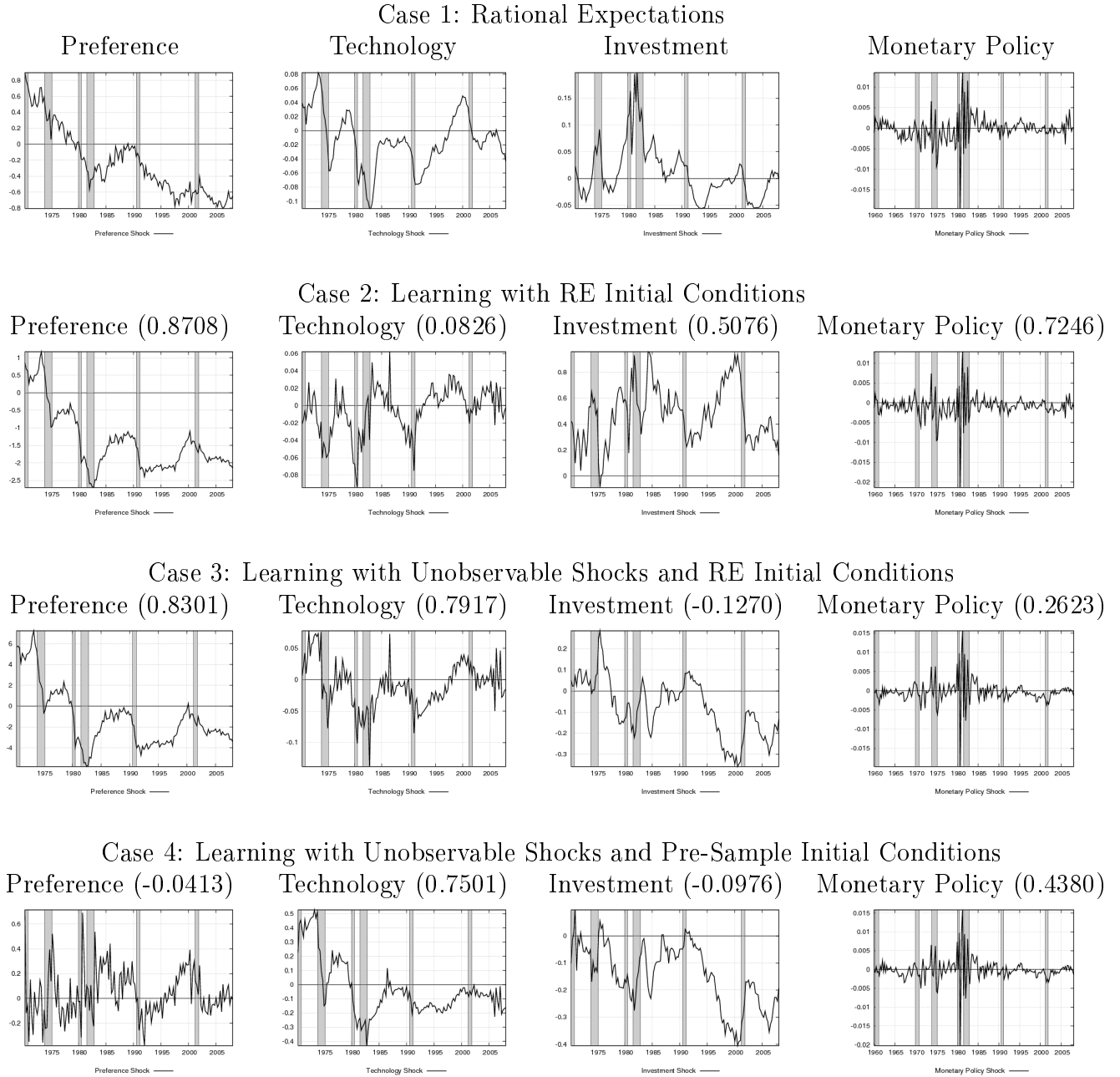


Table 2: Model Fit Comparisons

	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