Three Essays in Adaptive Expectations in New Keynesian Monetary Economies

James Murray

May 11, 2007

1. Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital
2. Regime Switching, Learning, and the Great Moderation
3. An Empirical Examination of Alternative Expectations Frameworks

- Purpose: Determine what features of U.S. data (if any) learning can explain.
- Learning has been suggested to deliver features:
  - Orphanides and Williams (2005): "Inflation scares".
  - Milani (2005): Persistence in output and inflation.
  - Volatility persistence of inflation.
  - Primiceri (2005): Great Moderation.
- Estimate a NK model with (constant gain) learning and RE by MLE.
- Examine forecast errors, evolution of shocks, and evolution of expectations.

Utility function:

$$U_0 = E_0^* \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-\sigma} \xi_t \left(c_t(i) - \eta c_{t-1}(i) \right)^{1-\sigma} - \frac{1}{1+\mu} n_t(i)^{1+\mu} \right]$$

- E_t^*: possibly non-rational expectations operator.
- $c_t(i)$: consumption at time t.
- $n_t(i)$: labor supply at time t.
- ξ_t: common preference shock.
- β: discount factor.
- $\sigma \in (0, \infty)$: related to the intertemporal elasticity of substitution.
- $\eta \in [0, 1)$: degree of habit formation.

Final good production:

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

- y_t output of final good, $y_t(i)$ output of intermediate good i.
- $\theta \in (1, \infty)$: elasticity of substitution in production.

Intermediate goods production:

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

- z_t: common technology shock.
- $k_t(i)$: firm-specific capital good.

- Final good is converted to a firm-specific capital good.
- Investment of $I_t(i)$ leads to capital stock next period:

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

- μ_t: common investment technology shock.
- δ: depreciation rate.
- ϕ: capital adjustment cost parameter.

- Follow Calvo (1983) pricing: fraction $1 - \omega$ firms re-optimize their price each period.
- Inflation indexation: Those who cannot re-optimize may adjust according to:

$$p_t(i) = p_{t-1}(i) + \gamma \pi_{t-1}$$

- Even with endogenous capital (Woodford, 2005), leads to Phillips curve:

$$\pi_t = \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{\beta}{1+\beta\gamma}E_t^*\pi_{t+1} + \frac{(1-\omega)(1-\omega\beta)}{\nu\omega(1+\beta\gamma)}\hat{s}_t$$

- $\hat{s}_t$: average marginal cost in the economy (percentage deviation from steady state).
- ν: function of many parameters (no closed form).

NK Model: Wrap up

- Monetary policy: $\hat{r}_t = \rho_r \hat{r}_{t-1} + (1-\rho_r)\left(\psi_\pi \pi_t + \psi_y \hat{y}_t\right) + \epsilon_{r,t}$
  - $\psi_\pi \in (0,\infty)$: feedback on inflation.
  - $\psi_y \in (0,\infty)$: feedback on output.
  - $\rho_r \in (0,1)$: smoothing parameter.
- Non-policy shocks (percentage deviations from steady state) are AR(1):
  - Preference shock: $\hat{\xi}_t = \rho_\xi \hat{\xi}_{t-1} + \epsilon_{\xi,t}$
  - Technology shock: $\hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t}$
  - Investment shock: $\hat{\mu}_t = \rho_\mu \hat{\mu}_{t-1} + \epsilon_{\mu,t}$
- Market clearing condition: $y_t = c_t + I_t$

Suppose a DSGE model of the form:

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

- x_t vector of time t variables, all observable to agents.
- E_t^*: possibly non-rational expectations operator.

Rational expectations solution implies:

$$E_t x_{t+1} = G x_t$$

- Agents know the form of this solution, but estimate elements of G by least squares.
- Use as explanatory variables past observations of x_t^k.
- x_t^k includes a constant and some or all variables in x_t.

- Let G_t^k include non-zero columns of G and a constant.
- Ordinary least squares estimate for G^k at time t:

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

- Least squares forecast:

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

- Evolution of $\hat{G}_t^k$ in recursive form:

$$\hat{G}_t^k = \hat{G}_{t-1}^k + g_t(x_{t-1} - \hat{G}_{t-1}^k x_{t-2}^k) x_{t-2}^k{}' R_t^{-1} \qquad (2)$$

$$R_t = R_{t-1} + g_t(x_{t-2}^k x_{t-2}^k{}' - R_{t-1}) \qquad (3)$$

- where $g_t = 1/(t-1)$ is the learning gain.

- Ordinary least squares $\rightarrow$ learning dynamics disappear.
- Constant gain:
  - Assumes g is constant.
  - Dynamics of expectations depend on the size of the learning gain.
  - Learning dynamics persist in the long run.
  - Appropriate (MSV solution) initial condition $+$ $(g = 0) \rightarrow$ RE.

Initializing Learning Algorithm

- $\hat{G}_t$ and R_t must be initialized for estimation.
- This paper: MSV solution, Expected Variance of state vector under RE.
  - Problem: Learning dynamics smallest when near RE solution.
- More popular: Use a VAR(1) from pre-sample data
  - Problem: initial condition far away from long run steady state.
  - Problem: latent variables - capital stock, structural shocks.
- Ad hock assumptions. In Milani (2005) coefficients on past inflation set to zero.
- Joint estimation: Zha (Lecture notes, 2005).
  - Problems: Possibly many additional parameters, possible over-fitting problem.

Estimation Procedure

- Estimate by MLE using Kalman filter.
- Data: Quarterly data from 1957 through 2005.
  - Output growth: growth rate of real GDP.
  - Investment growth: growth rate of real aggregate domestic investment.
  - Inflation: growth rate of GDP deflator.
  - Interest rate: federal funds rate.

Results with No Capital Accumulation

Description	Parameter	Learning	RE
Learning gain	g	0.017225***	–
Discount factor	β	0.994528***	0.994893***
Habit formation	η	0.269967**	0.213315
Inverse elasticity sub.	σ	16.264823	17.589033
Elasticity of sub. production	θ	11.459122	10.030577
Inverse elasticity labor supply	μ	2.509288	1.791766
Calvo parameter	ω	0.715756	0.713950
Inflation indexation	γ	0.441247***	0.958006***
MP interest rate smoothing	ρ_r	0.866560***	0.823025***
MP feedback on output	ψ_y	0.123166**	0.055319*
MP feedback on inflation	ψ_π	0.995049***	1.067211***
Pref. shock persistence	ρ_ξ	0.931082***	0.962095***
Tech. shock persistence	ρ_z	0.000010	0.000010
Steady state inflation	π^*	2.814514***	3.296221**
Std. dev. technology shock	σ_z	0.301741*	0.340167
Std. dev. preference shock	σ_ξ	0.280402*	0.254504
Std. dev. interest rate shock	σ_r	0.002283***	0.002344***

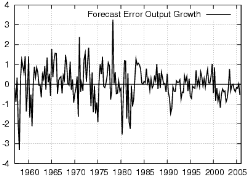
* Significantly different from zero at the 10% level.
** Significantly different from zero at the 5% level.
*** Significantly different from zero at the 1% level.

- Learning statistically significant.
- Learning leads to lower degree of inflation indexation.
- Habit formation still significant source of persistence.
- Very similar estimates for the degree of price flexibility.
- Possible reasons for differences with Milani (2005):
  - MLE vs. Bayesian methods.
  - Different initial condition for recursive learning process.
  - Different data: Growth rate of real GDP vs. CBO measure of output gap.

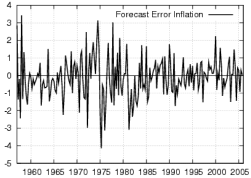
In-Sample One Quarter Ahead Forecast Errors

Learning

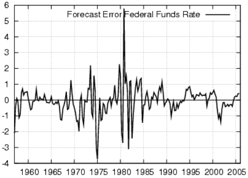
Output



Inflation

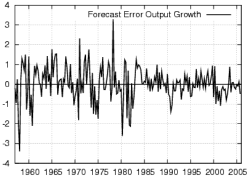


Interest Rate

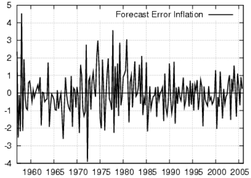


Rational Expectations

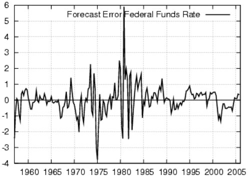
Output



Inflation



Interest Rate

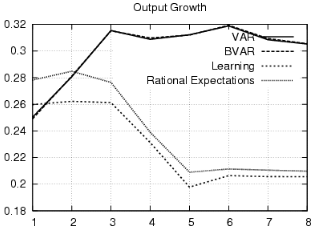


- Forecast errors are very similar for Learning and RE.
- Forecast errors are more volatile in 1970s.
- Huge forecast errors for federal funds rate in late 70s, early 80s.
- Learning appears not to be explaining U.S. dynamics better than RE.
- Both fail to explain high volatility in 70s with low volatility after mid 80s.

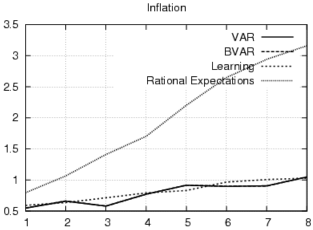
Eight Quarter Out-of-Sample Forecasts

- Estimate the model through 1989:Q4.
- Use estimated parameters to forecast 1990:Q1 - 2005:Q4.
- For each quarter, forecast eight periods ahead.
- Given forecasts of each horizon, compute MSE.
- Do the same for a VAR(4) and Litterman (1986) BVAR(4).

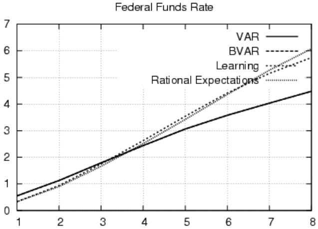
Output



Inflation



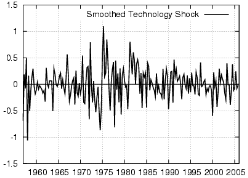
Interest Rate



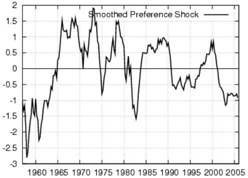
Smoothed Technology Shocks

Learning

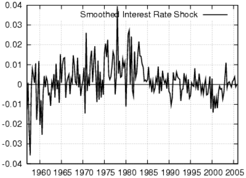
Technology Shock



Preference Shock

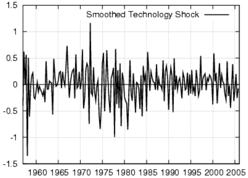


Interest Rate Shock

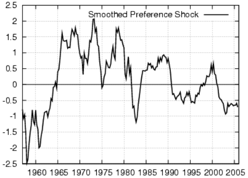


Rational Expectations

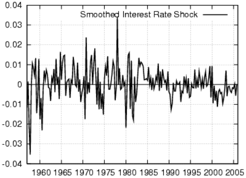
Technology Shock



Preference Shock



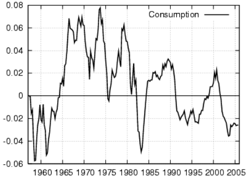
Interest Rate Shock



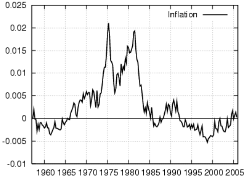
Agents' Expectations

Learning

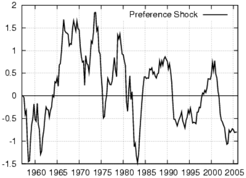
Consumption



Inflation

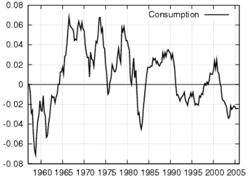


Preference shock

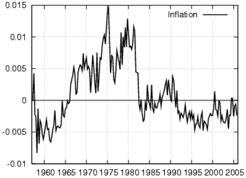


Rational Expectations

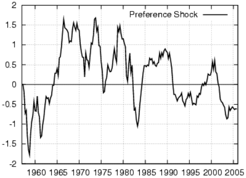
Consumption



Inflation



Preference shock



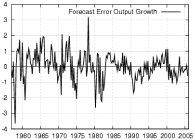
Description	Parameter	Learning	RE
Learning gain	g	0.024327**	–
Discount factor	β	0.993366***	0.992763***
Habit formation	η	0.281563**	0.286030*
Inverse elasticity sub.	σ	18.558950	16.328695
Elasticity of sub. production	θ	13.122359	7.317025
Inverse elasticity labor supply	μ	3.329474	6.219549
Capital share of income	α	0.174750	0.186636
Depreciation rate	δ	0.163489	0.299872
Cost of adjusting capital	ϕ	13.455016	11.558055
Calvo parameter	ω	0.658099	0.774919
Inflation indexation	γ	0.404221***	0.760702***
MP interest rate smoothing	ρ_r	0.869496***	0.859279***
MP feedback on output	ψ_y	0.064696*	0.128857***
MP feedback on inflation	ψ_π	0.992672***	0.967512***
Pref. shock persistence	ρ_ξ	0.984689***	0.980813***
Tech. shock persistence	ρ_z	0.012960	0.000010
Inv. shock persistence	ρ_μ	0.804935***	0.824007***
Steady state inflation	π^*	3.570552***	4.073905***
Std. dev. technology shock	σ_z	0.229175**	0.400000**
Std. dev. investment shock	σ_μ	0.060246***	0.052502***
Std. dev. preference shock	σ_ξ	0.231848*	0.214337*
Std. dev. interest rate shock	σ_r	0.002291***	0.002286***

- Learning remains statistically significant.
- Including capital leads to a lower degree of inflation indexation.
- Learning further lowers the degree of inflation indexation.
- Habit formation still significant source of persistence.
- Learning leads to lower estimates for the degree of price flexibility.
- Learning leads to lower estimates for the variance of the technology shock.

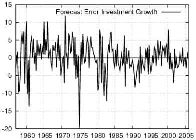
One Quarter Ahead Forecast Errors

Learning

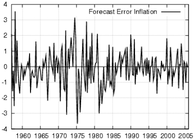
Output



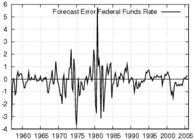
Investment



Inflation

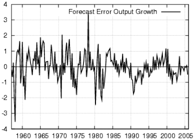


Interest Rate

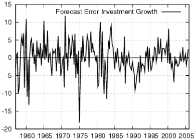


Rational Expectations

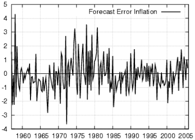
Output



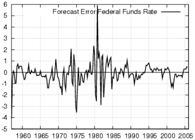
Investment



Inflation

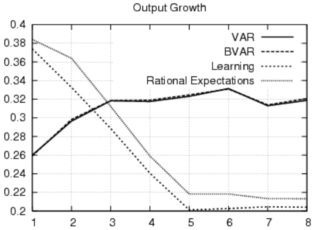


Interest Rate

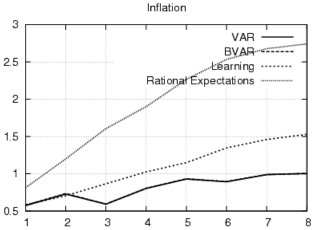


Eight Quarter Out-of-Sample Forecasts

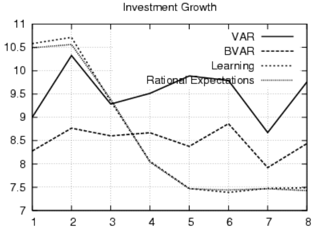
Output



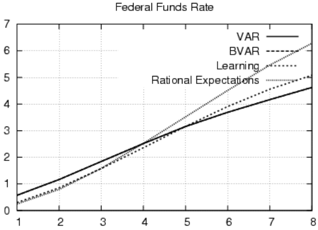
Inflation



Investment

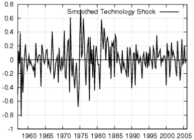
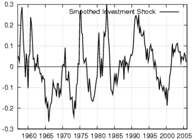
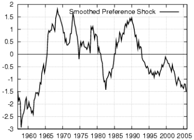
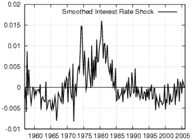


Interest Rate

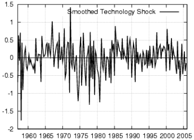
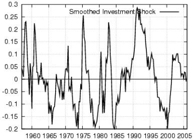
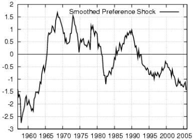
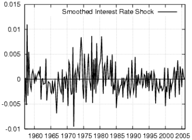


Smoothed Technology Shocks

Learning

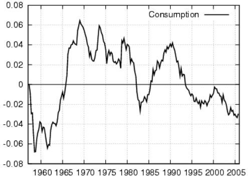
Technology Shock	Investment Shock	Preference Shock	Interest Rate Shock
			

Rational Expectations

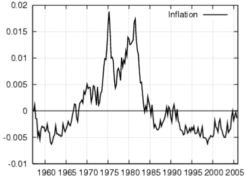
Technology Shock	Investment Shock	Preference Shock	Interest Rate Shock
			

Learning

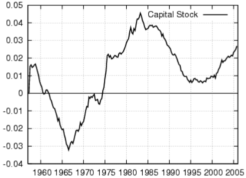
Consumption



Inflation

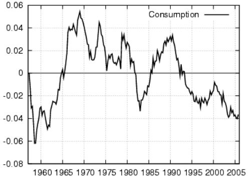


Capital stock

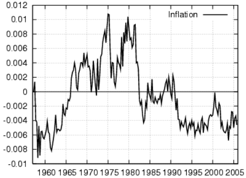


Rational Expectations

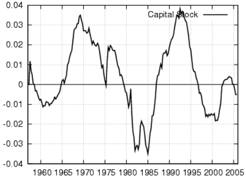
Consumption



Inflation



Capital stock



Conclusion

- Learning (small) successes:
  - Lower estimates for inflation indexation, Calvo parameter (using capital).
  - Some evidence of inflation scares.
  - Better performance out-of-sample, especially for inflation.
- Learning failures:
  - Very similar to RE in explaining the data.
  - 1970s inflation, monetary policy.
- Hypotheses:
  1. Sensitivity to initial conditions suggests learning explains better dynamics following structural change.
  2. Dynamic gain may deliver larger effects.
  3. Williams (2003), Primiceri (2005): Learning about structural features may deliver quantitatively larger effects.

- Addresses my first two hypotheses.
- "Bad luck" regime changes (not my language).
  - Variances of non-policy shock switch between high/low states according to Markov chain.
  - Sims and Zha (2006): evidence of Great Moderation points towards this type of switching.
  - Bullard and Singh (2007): bad luck + Bayesian learning can explain Great Moderation.
- Learning framework: Marcet and Nicolini (2003) endogenous learning gain.

- When recent forecast errors get big, agents switch to higher learning gain (higher discount for past data).
- Let $\alpha_t \equiv 1/g_t$. In OLS case α_t is sample size.

$$\alpha_t = \begin{cases} \alpha_{t-1} + 1 & \text{if } \frac{1}{J}\sum_{j=1}^{J}\frac{1}{n}\sum_{v=1}^{n}\left|\frac{x_{t-j}(v) - \hat{G}^*_{t-j}(v)x^*_{t-j-1}}{\hat{G}^*_{t-j}(v)x^*_{t-j-1}}\right| < \nu \\ \alpha & \text{otherwise} \end{cases} \tag{4}$$

- Notation:
  - n is the number of variables in x_t.
  - $x_{t-j}(v)$: vth element of x_{t-j}.
  - $\hat{G}^*_{t-j}(v)$: vth row of $\hat{G}^*_{t-j}(v)$.
  - $\alpha \equiv 1/g$ is the constant gain.
  - $\nu \in (0, \infty)$ is a threshold level.

- Serially correlated shocks:

$$z_t = \rho_z z_{t-1} + \epsilon_{z,t}(s_t) \tag{5}$$

$$\xi_t = \rho_\xi \xi_{t-1} + \epsilon_{\xi,t}(s_t) \tag{6}$$

- Innovations for a given regime are mean zero and iid.

$$Var\left(\epsilon_{i,t}(s_t)\right) = \begin{cases} \sigma^2_{i,1}, \text{ if } s_t = 1 \\ \sigma^2_{i,1}, \text{ if } s_t = 2 \end{cases} \tag{7}$$

- Where $i = \xi, z, \mu$ (preference, technology, and investment shocks).
- Transition matrix:

$$P = \begin{bmatrix} p_1 & 1 - p_1 \\ 1 - p_2 & p_2 \end{bmatrix}. \tag{8}$$

- Kim (1994), Kim and Nelson (1999): form the likelihood.
  - Kalman filter for state-space models.
  - Hamilton (1989) filter for regime switching.
- Plan to estimate:
  - Point estimates of high and low variances: $\sigma^2_{z,1}, \sigma^2_{\xi,1}$, and $\sigma^2_{z,1}, \sigma^2_{\xi,1}$.
  - Learning gain parameters: g and ν.
  - Smoothed estimates of the regime probabilities for each period.
  - Smoothed estimates of the shocks.
  - Out-of-sample forecasts of the model.

Interesting questions

- Can one identify regime switches and dynamic learning gain changes simultaneously?
- If so, does learning play a larger role with regime changes?
- With endogenous gain learning, how sizable of an increase in shock volatility is necessary to explain volatile periods such as the 1970s.

- Deviating from rational expectations is becoming more popular.
- Alternative frameworks even in least squares learning:
    - Decreasing, constant, dynamic gain.
    - Structural learning.
    - Preston (2005): Long run horizon learning.
    - Branch, Carlson, Evans, and McGough (2006a, 2006b): Endogenous inattention.
- Outside of least squares learning:
    - Bayesian learning.
    - Rational inattention.

Goal of this paper

- Estimate same NK model under alternative expectations frameworks.
- Examine forecast errors, smoothed shocks, agents expectations.
- Generate impulse response functions (IRF) on estimated model.
- Interesting questions:
    - Are some expectations frameworks easier to identify?
    - How well to different expectations frameworks explain U.S. experience of changing volatility and changing persistence better?
    - Does the expectations framework matter? They are not mathematically equivalent; but are they observationally equivalent?
    - Do IRF from each framework imply different dynamics?