Discussion of:

Expectations, Learning, and Macroeconomic Persistence

by Fabio Milani

Eric Swanson

Federal Reserve Bank of San Francisco

Conference on DSGE Models and Macroeconomic Forecasting

Federal Reserve Bank of Cleveland

October 14, 2005

Summary of Paper

1. Lucas-style critique of “hybridized” New Keynesian models

2. Joint Bayesian estimation of *all* structural parameters, including learning

3. Learning is sufficient to explain macro persistence

Lucas-Style Critique of Hybridized NK Models

$$\pi_t = \mu_1 E_t \pi_{t+1} + (1-\mu_1)\pi_{t-1} + \gamma y_t + \varepsilon_t^{\pi}$$

$$y_t = \mu_2 E_t y_{t+1} + (1-\mu_2) y_{t-1} - \beta(i_t - E_t \pi_{t+1}) + \varepsilon_t^{y}$$

Presence of lags justified by:
- consumer habits
- indexation to previous period’s inflation rate
- rule-of-thumb consumers or firms
- higher-order adjustment costs

Little micro evidence in support

Alternative source of persistence: learning

Lucas-Style Critique of Hybridized NK Models

In general, structural source of persistence matters:

- Welfare (e.g., costs of inflation)
- Optimal policy
- Cross-equation restrictions
- Counterfactuals
- Structural breaks (e.g., in inflation persistence)
- Forecasting
- Simulations without historical precedent, or with bad historical precedent (e.g., oil price shocks)

of course, this critique applies to some other aspects of the model as well (e.g., Calvo contracts)

Methodology

Learning:

- Agent and Firm expectations formed from a VAR (essentially)
- Coefficients of VAR updated via constant-gain learning

Previous work on learning:

- E-stability
- Calibration
- Estimation of gain coefficient from survey data
  (Branch-Evans, Orphanides-Williams)

In this paper:

- Full system estimation of *all* parameters of model, including learning
- Bayesian methods

Description	Parameters	Range	Prior Distr.	Prior Mean	Prior Std.	95% Prior Prob. Int.
Habit Formation	η	$[0, 1]$	*Uniform*	.5	.289	$[0.025, 0.975]$
Discount rate	β	$[0, 1]$	*Beta*	.99	.01	$[0.973, 0.999]$
IES	σ	$\mathbb{R}^+$	*Gamma*	.125	.09	$[0.015, 0.35]$
Infl. Indexation	γ	$[0, 1]$	*Uniform*	.5	.289	$[0.025, 0.975]$
Function price stick.	ξ_p	$\mathbb{R}$	*Gamma*	.015	.011	$[0.0019, 0.04]$
Elast. mc to inc.	ω	$\mathbb{R}$	*Normal*	.8975	.4	$[0.114, 1.68]$
Int-rate smooth.	ρ	$[0, 0.97]$	*Uniform*	.485	.28	$[0.024, 0.946]$
Feedback Infl.	χ_π	$\mathbb{R}$	*Normal*	1.5	.25	$[1.01, 1.99]$
Feedback Gap	χ_x	$\mathbb{R}$	*Normal*	.5	.25	$[0.01, 0.99]$
Autoregr. Dem shock	ϕ_r	$[0, 0.97]$	*Uniform*	.485	.28	$[0.024, 0.946]$
Autoregr. Sup shock	ϕ_u	$[0, 0.97]$	*Uniform*	.485	.28	$[0.024, 0.946]$
MP shock	σ_ε	$\mathbb{R}^+$	*InvGamma*	1	.5	$[0.34, 2.81]$
Demand shock	σ_r	$\mathbb{R}^+$	*InvGamma*	1	.5	$[0.34, 2.81]$
Supply shock	σ_u	$\mathbb{R}^+$	*InvGamma*	1	.5	$[0.34, 2.81]$
Gain Coeff.	$\overline{\mathbf{g}}$	$\mathbb{R}^+$	*Gamma*	.031	.022	$[0.0038, 0.087]$

Table 1 - Prior distributions for model with learning.

Notes: The same priors are used for the model with rational expectations, excluding the prior for the gain coefficient $\overline{\mathbf{g}}$, not needed.

Results: Learning is Sufficient

Nested model analysis:

- Estimate the model with learning, habits, indexation
- Learning is highly significant—RE is rejected
- Coefficients on habits, indexation are zero

Non-nested model analysis:

- Model with only learning is vastly favored over habits & indexation

	Learning	Rational Expectations
Log Marginal Likelihood	-759.08 (1.326)	-765.45 (1.316)
Posterior odds	584.06	1
Posterior probabilities	0.99829	0.0017092

Table 4 - Model Comparison: Learning (with frictions) vs. Rational Expectations (with frictions).

	Learning	Rational Expectations
Log Marginal Likelihood	-750.65 (1.375)	-765.45 (1.316)
Posterior odds	$2.6764 * 10^{6}$	1
Posterior probabilities	1	$3.7363 * 10^{-7}$

Table 5 - Model Comparison: Learning (no frictions) vs. Rational Expectations (with frictions).

Comment 1: Take Structural Critique More Seriously

Learning in the model is reduced-form ("Euler equation" learning):

$$\pi_t = \mu_1 \hat{E}_t \pi_{t+1} + (1-\mu_1)\pi_{t-1} + \gamma y_t + \varepsilon_t^\pi$$

$$y_t = \mu_2 \hat{E}_t y_{t+1} + (1-\mu_2) y_{t-1} - \beta(i_t - E_t \pi_{t+1}) + \varepsilon_t^y$$

Better analysis would be structural ("infinite-horizon" learning, Preston):

$$\pi_t = \hat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left[\xi_p \omega y_T + (1-\alpha)\beta\pi_{T+1} + \varepsilon_T^\pi \right]$$

$$y_t = \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \left[(1-\beta) y_{T+1} - (1-\eta\beta)\sigma\left(i_T - \pi_{T+1}\right) + \varepsilon_T^\pi \right]$$

Results:

- Calvo price adjustment more frequent (duration of 1.2 qtrs)
- Autocorrelation of supply shock lower (0.2)
- Fit of model is worse

Parameters	Mean estimate	95% Post. Prob. Interval
η	0.171	$[0.04, 0.32]$
β	0.989	$[0.97, 0.999]$
α	0.138	$[0.02, 0.36]$
σ	0.051	$[0.02, 0.10]$
γ	0.204	$[0, 0.78]$
ξ_p	0.0015	–
ω	0.8975	–
ρ	0.894	$[0.85, 0.93]$
χ_π	1.508	$[1.14, 1.84]$
χ_x	0.797	$[0.41, 1.25]$
ϕ_r	0.902	$[0.82, 0.96]$
ϕ_u	0.017	$[0, 0.06]$
σ_ε	0.934	$[0.84, 1.04]$
σ_r	0.822	$[0.74, 0.91]$
σ_u	1.382	$[1.12, 2.04]$
$\overline{\mathbf{g}}$	0.0283	$[0.009, 0.048]$

Table 6 - Infinite-Horizon Learning: posterior estimates and 95% posterior probability intervals.

Comment 2: What About Other Sources of Persistence?

Sticky Wages
Levin-Wolman contracts
Capital

Habit persistence in π^* (as opposed to simple random walk)
(e.g., Gurkaynak, Sack, Swanson)

Comment 3: Consider Richer Learning Specification

Agents' VARs are based on MSV set
Note: this differs across models
    Model with habits, indexation has more state variables

How do agents know which model they're in?

Why not base agents' expectations on a richer set of variables?

Presumably, this would make learning more difficult, increase persistence

Comment 4: Estimation Bells & Whistles

Estimate potential output y*

Estimate model using survey (SPF) data as well as actual data

Another Source of Persistence: Central Bank Habit Persistence (in π^*)

$$\pi_t = \mu_1 E_t \pi_{t+1} + (1-\mu_1)\pi_{t-1} + \gamma y_t + \varepsilon_t^{\pi}$$

$$y_t = \mu_2 E_t y_{t+1} + (1-\mu_2) y_{t-1} - \beta(i_t - E_t \pi_{t+1}) + \varepsilon_t^{y}$$

$$i_t = (1-c)\left[\bar{\pi}_t + a(\bar{\pi}_t - \pi_t^*) + by_t\right] + c\, i_{t-1} + \varepsilon_t^{i}$$

$$\pi_t^* = \pi_{t-1}^* + \theta\left(\bar{\pi}_{t-1} - \pi_{t-1}^*\right) + \varepsilon_t^{\pi^*}$$

This model fits the U.S. term structure response to economic news very well
(fixed π^* fails to capture responsiveness of long-term nominal rates)

Question: how do central bank habits in π^* affect:

Short-run model dynamics

Inflation persistence

Estimates of μ_1 and μ_2

Central Bank Habit Persistence (in π*)

Start by estimating standard model:

$$\pi_t = \mu_1 E_t \pi_{t+1} + (1-\mu_1)\pi_{t-1} + \gamma y_t + \varepsilon_t^{\pi}$$

$$y_t = \mu_2 E_t y_{t+1} + (1-\mu_2) y_{t-1} - \beta(i_t - E_t \pi_{t+1}) + \varepsilon_t^{y}$$

$$i_t = (1-c)\left[\bar{\pi}_t + a(\bar{\pi}_t - \pi_t^{*}) + b y_t\right] + c\, i_{t-1} + \varepsilon_t^{i}$$

Using quarterly data on π, y, i
9 parameters, 240 observations (using post-1985 period)

Let's use full information (system) Maximum Likelihood
with parameter constraints (Matlab fmincon)
imposing Blanchard-Kahn stability or near-stability (i.e., penalize instability)

```
Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
No active inequalities
Aim: too many big roots.
  242.95  MU=[0.153 0.462] AY=0.03 GY=[1.08 0.24] GR=0.00 RHO=1.04 MP=[3.35 3.82]


Optimization terminated: magnitude of directional derivative in search
 direction less than 2*options.TolFun and maximum constraint violation
  is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    6          4
    8
  199.84  MU=[0.620 0.386] AY=0.02 GY=[1.20 -0.21] GR=0.00 RHO=0.95 MP=[0.50 1.43]
```

relax GY constraint to GY(1)<1.4 (was 1.2):

```
Optimization terminated: magnitude of directional derivative in search
 direction less than 2*options.TolFun and maximum constraint violation
  is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    6
    8
  199.84  MU=[0.620 0.002] AY=0.02 GY=[1.20 -0.24] GR=0.00 RHO=0.95 MP=[0.50 1.42]


Optimization terminated: magnitude of directional derivative in search
 direction less than 2*options.TolFun and maximum constraint violation
  is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
```

```
    6
    8
```

relax MP(1) constraint to MP(1)>0.3:

```
Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
No active inequalities
  199.95  MU=[0.569 0.477] AY=0.03 GY=[1.25 -0.24] GR=0.00 RHO=1.05 MP=[3.05 0.50]
```

relax constraint on MP to MP(2)>0:

```
Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
No active inequalities
Aim: too many big roots.
  240.45  MU=[0.140 0.504] AY=0.02 GY=[1.11 0.46] GR=0.12 RHO=0.88 MP=[1.06 1.14]
```

```
Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    6
  205.38  MU=[0.400 0.419] AY=0.05 GY=[1.20 -0.20] GR=0.00 RHO=1.00 MP=[0.95 2.46]
```

```
Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
No active inequalities
  918.45  MU=[0.769 0.052] AY=0.27 GY=[0.47 -0.15] GR=0.13 RHO=1.15 MP=[0.90 0.76]
```

```
Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    6          1
               4
  574.35  MU=[0.000 0.066] AY=0.00 GY=[0.68 -0.19] GR=-0.00 RHO=0.75 MP=[1.91 1.81]


Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    2
    6
    9
 1175.74  MU=[0.957 0.000] AY=2.01 GY=[0.56 -0.28] GR=0.00 RHO=0.99 MP=[4.35 -0.00]


Optimization terminated: magnitude of directional derivative in search
 direction less than 2*options.TolFun and maximum constraint violation
  is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    6
    8
  199.71  MU=[0.620 0.049] AY=0.02 GY=[1.20 -0.24] GR=0.00 RHO=0.95 MP=[0.30 1.43]


Optimization terminated: magnitude of search direction less than 2*options.TolX
 and maximum constraint violation is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
```

```
    6
  199.98  MU=[0.619 0.433] AY=0.02 GY=[1.20 -0.20] GR=0.00 RHO=0.96 MP=[0.51 1.38]


Optimization terminated: magnitude of directional derivative in search
 direction less than 2*options.TolFun and maximum constraint violation
  is less than options.TolCon.
Active inequalities (to within options.TolCon = 1e-006):
  lower      upper     ineqlin   ineqnonlin
    6
    8
  199.71  MU=[0.620 0.124] AY=0.02 GY=[1.20 -0.24] GR=0.00 RHO=0.95 MP=[0.30 1.43]
```

Summary:

- Likelihood surface looks like an egg carton, or worse
- μ_2 in particular is not well-identified by the data
- IS curve response to interest rates estimated to be 0

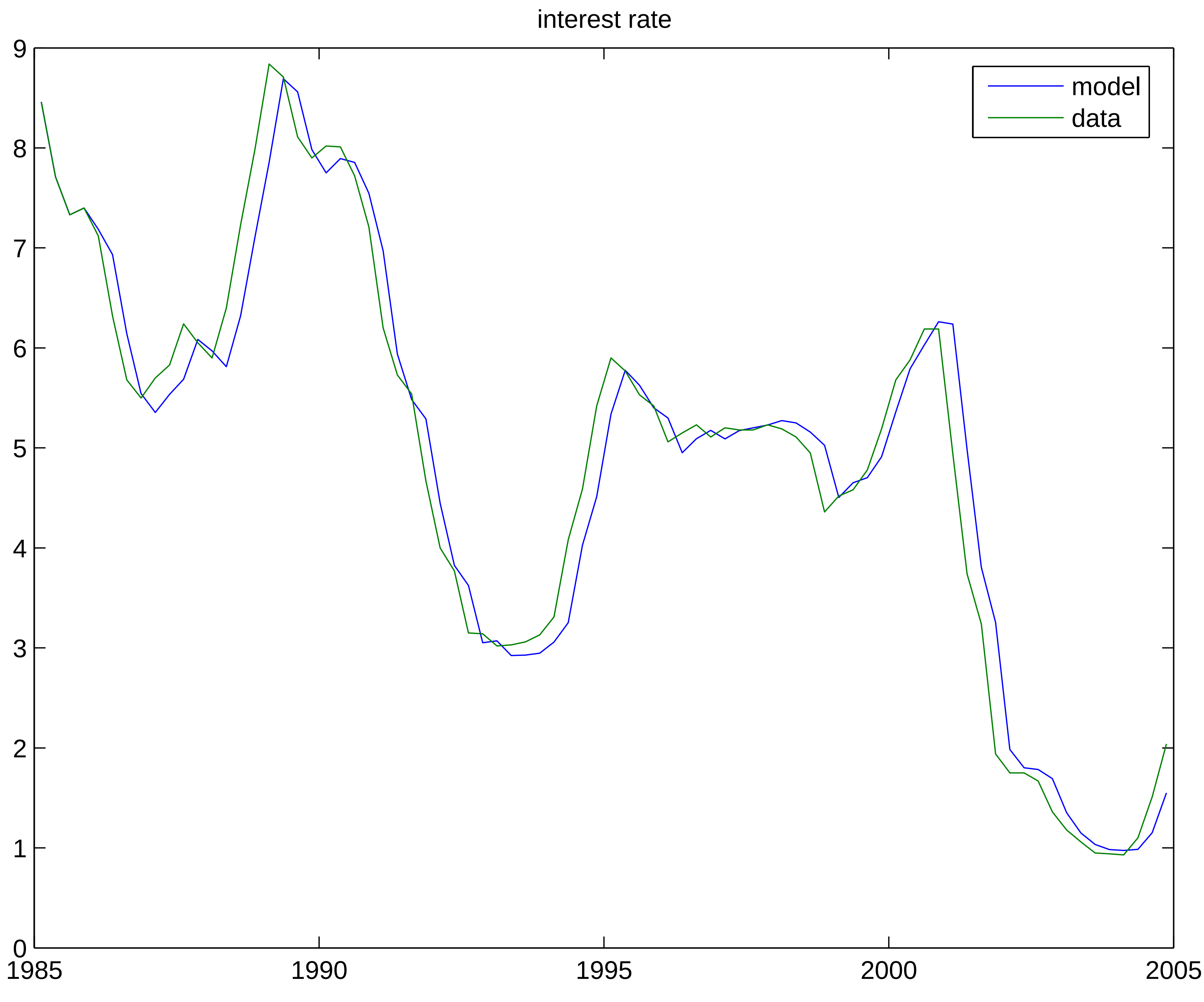
Is Bayesian estimation the answer?

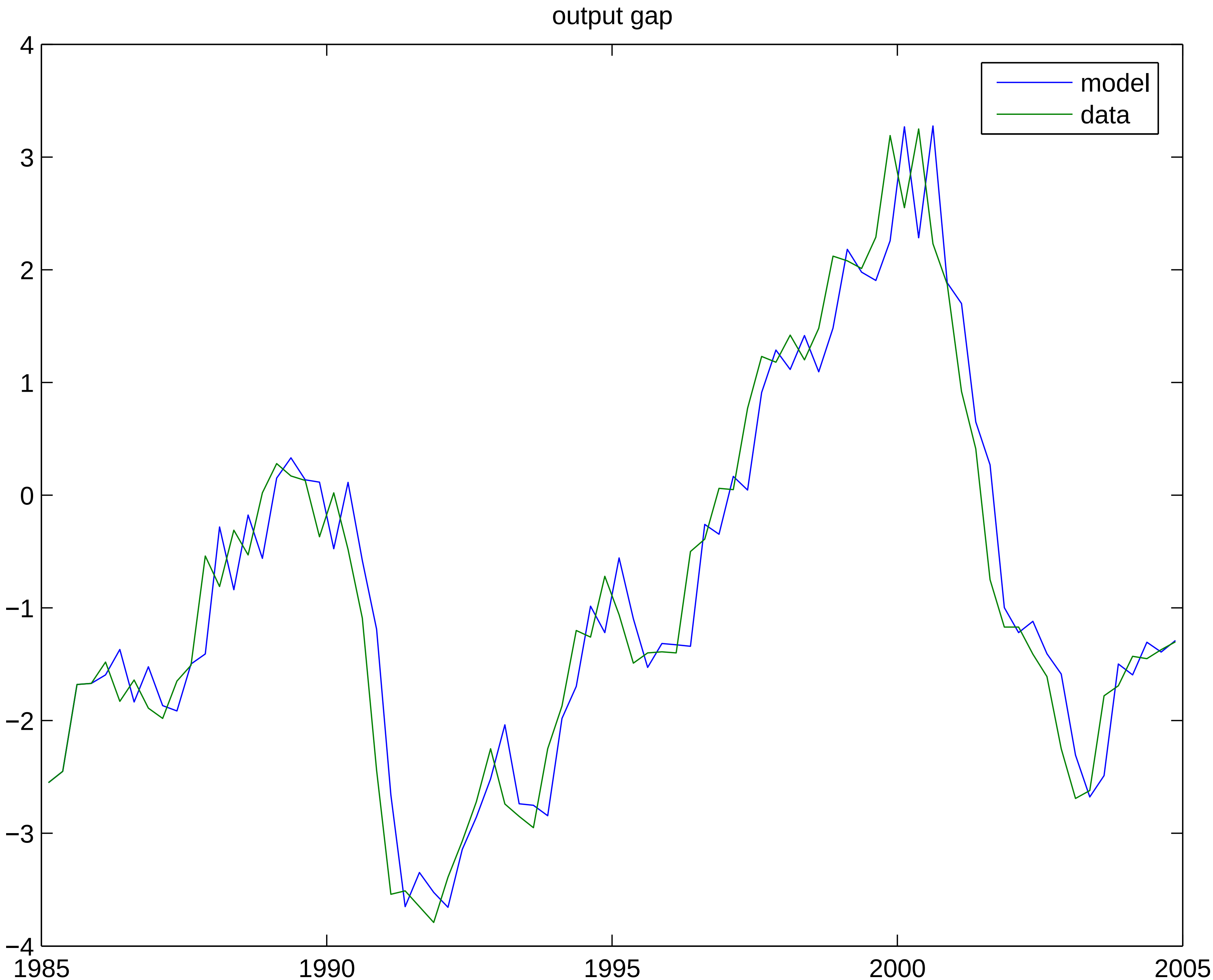
- Priors mix additional curvature into likelihood function
- Using posterior *mean* (instead of mode) stabilizes the estimate
- But fundamental lack of identification is still present
  - should be reflected in the posterior confidence intervals

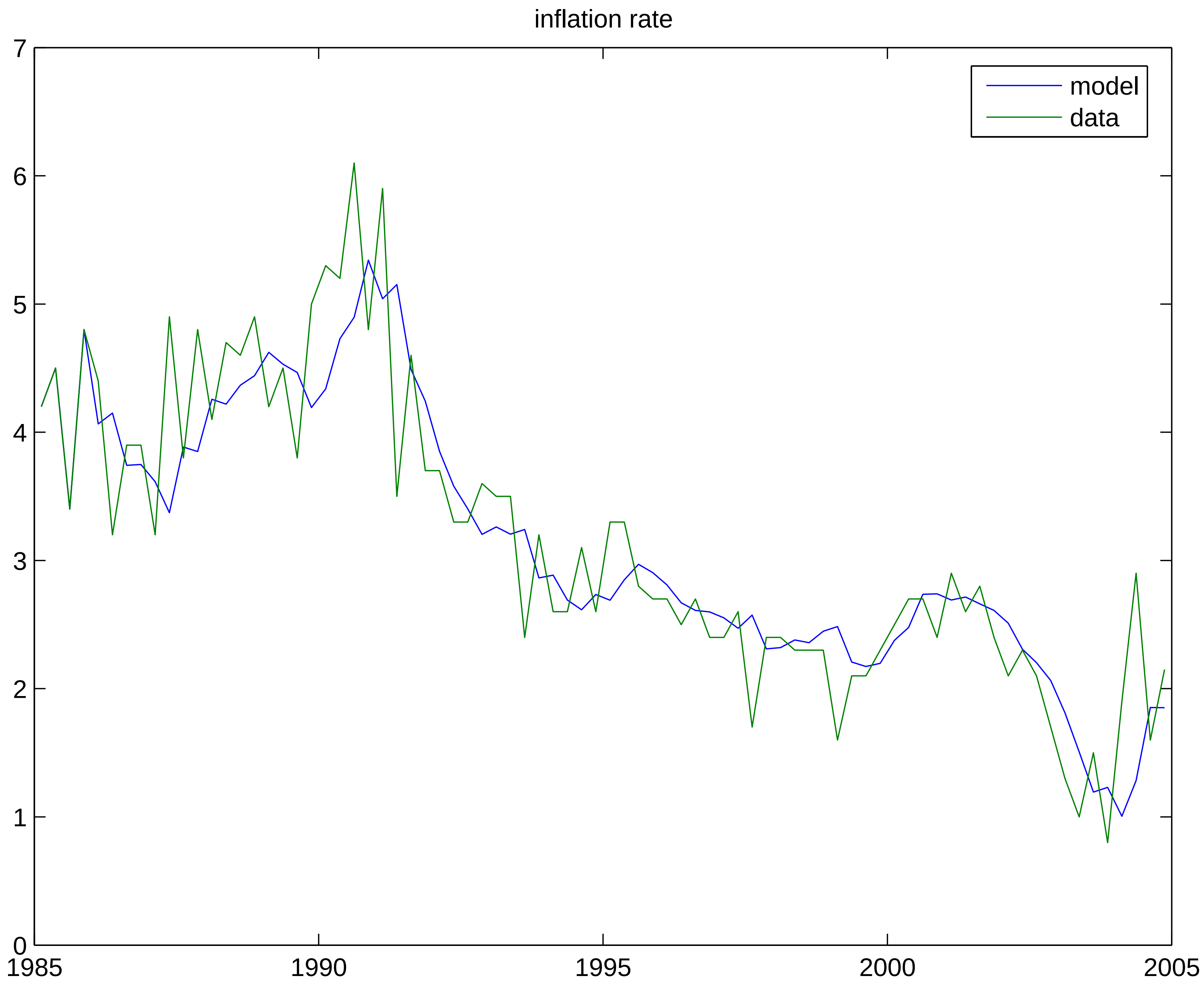
Description	Parameters	Giannoni-Woodford '03	Bayesian Estimation	
		Estimate	Mean Estimate	95% Post. Prob. Int.
Habits	η	1	0.911	$[0.717, 0.998]$
Discount	β	0.99 (fixed)	0.9897	$[0.971, 0.999]$
IES	φ	0.6643	3.813	$[2.285, 6.02]$
Indexation	γ	1	0.885	$[0.812, 0.957]$
Fcn. price stick.	ξ_p	0.0015	0.001	$[0.0001, 0.002]$
Elast. mc	ω	0.8975	0.837	$[0.01, 1.63]$
Int-rate smooth.	ρ	–	0.89	$[0.849, 0.93]$
Feedback Infl.	χ_π	–	1.433	$[1.06, 1.81]$
Feedback Gap	χ_x	–	0.792	$[0.425, 1.165]$
Autoregr. Dem shock	ϕ_r	–	0.87	$[0.8, 0.93]$
Autoregr. Sup shock	ϕ_u	–	0.02	$[0.0005, 0.07]$
MP shock	σ_ε	–	0.933	$[0.84, 1.04]$
Demand shock	σ_r	–	1.067	$[0.89, 1.22]$
Supply shock	σ_u	–	1.146	$[1.027, 1.27]$

Table 3 - Rational Expectations Estimates and 95% posterior probability interval.

Note: 0.0187% of the draws fell in the indeterminacy region and were discarded.


interest rate
model
data
9
8
7
6
5
4
3
2
1
0
1985
1990
1995
2000
2005


output gap
model
data
4
3
2
1
0
−1
−2
−3
−4
1985
1990
1995
2000
2005



inflation rate
model
data
7
6
5
4
3
2
1
0
1985
1990
1995
2000
2005

Conclusions

Paper is very thorough, well-written, very nicely done

Lots of robustness checks and sensitivity analysis

Raises good questions and offers answers:
Do we need ad hoc persistence in the NK model?
Learning: a unified formulation of persistence?

Lots of potential modifications for future research:
Wage contracts, Levin-Wolman contracts, capital, Smets-Wouters and CEE models, y* estimation, central bank habit persistence

Bayesian estimation: too good to be true?