Leveling the Playing Field

Prior Choice and DSGE Model Comparisons

Marco Del Negro

Federal Reserve Bank of Atlanta Fed

Frank Schorfheide

University of Pennsylvania, CEPR

Practical Issues in DSGE Modelling at Central Banks
Bank of Finland, June 2006

Disclaimer: The views expressed are the author's and do not necessarily reflect those of the Federal Reserve Bank of Atlanta or the Federal Reserve System

Motivation

- The new generation of new-Keynesian DSGE models (Christiano et al., Smets and Wouters ...) fits the data reasonably well, and hence can be used for policy analysis at Central Banks.

- These models contain many bells and whistles (and persistent shocks) – some are more “structural” than others.

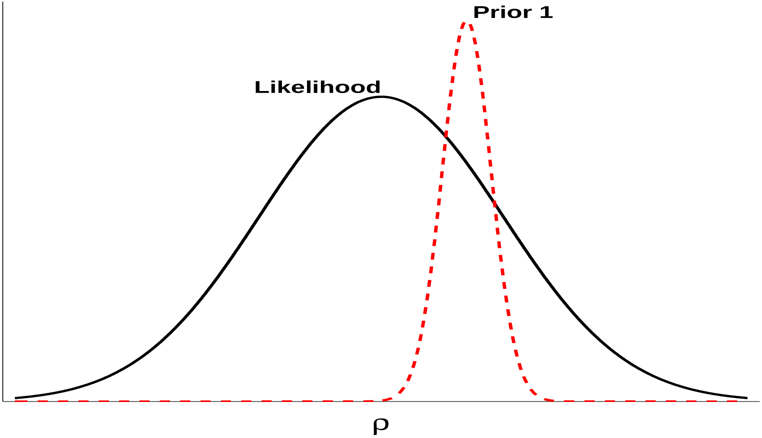
- Which features are really needed, and which can we get rid of?

Motivation

- The new generation of new-Keynesian DSGE models (Christiano et al., Smets and Wouters ...) fits the data reasonably well, and hence can be used for policy analysis at Central Banks.
- These models contain many bells and whistles (and persistent shocks) – some are more "structural" than others.
- Which features are really needed, and which can we get rid of?
- Two approaches for model comparison:
  - Impulse responses (CEE)
  - Bayesian model comparisons via Marginal Likelihoods (Smets and Wouters)

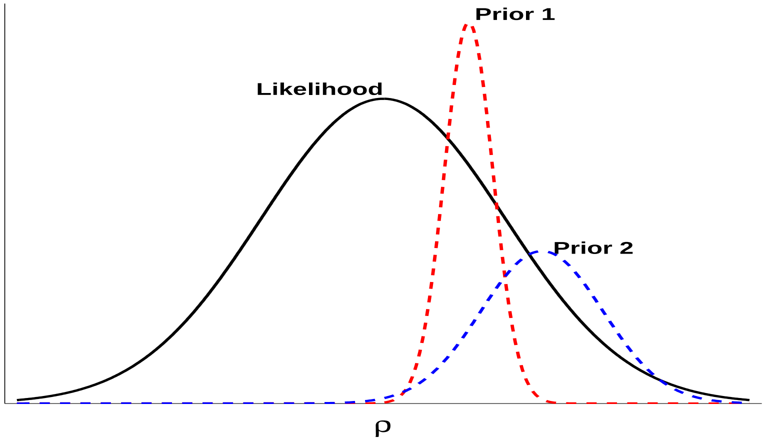
Priors and Model Comparisons

- The marginal likelihood is the integral of the likelihood with respect to the prior



Priors and Model Comparisons

- The marginal likelihood is the integral of the likelihood with respect to the prior
- . . . hence the choice of the prior matters



Leveling the Playing Field

- In Bayesian model comparisons, priors should be chosen so that all models are given a fair chance "a priori".

Leveling the Playing Field

- In Bayesian model comparisons, priors should be chosen so that all models are given a fair chance "a priori".
- We focus on priors for the auxiliary parameters (correlation and st. dev. of exogenous shocks):
  - Hard to have intrinsic beliefs about the driving process of these unobservable shocks
  - ... but we do have beliefs about the implications for the observables (i.e., volatility of inflation, etc.).

Leveling the Playing Field

- In Bayesian model comparisons, priors should be chosen so that all models are given a fair chance "a priori".

- We focus on priors for the auxiliary parameters (correlation and st. dev. of exogenous shocks):

  - Hard to have intrinsic beliefs about the driving process of these unobservable shocks

  - . . . but we do have beliefs about the implications for the observables (i.e., volatility of inflation, etc.).

1. Choose priors so that the implications for the endogenous variables are close across models.

2. Introduce dependence among parameters.

Identifying Backward Looking Behaviour in a Simple Example

- Take two models:

$$\mathcal{M}_1: \quad y_t = \frac{1}{\alpha} \boldsymbol{E}_t[y_{t+1}] + \rho y_{t-1} + u_t, \quad u_t = \epsilon_t \sim iid(0, \sigma^2).$$

$$\mathcal{M}_2: \quad y_t = \frac{1}{\alpha} \boldsymbol{E}_t[y_{t+1}] + u_t, \quad u_t = \rho u_{t-1} + \epsilon_t \sim iid(0, \sigma^2).$$

Identifying Backward Looking Behaviour in a Simple Example

- Take two models:

$$\mathcal{M}_1 : \quad y_t = \frac{1}{\alpha}\boldsymbol{E}_t[y_{t+1}] + \rho y_{t-1} + u_t, \quad u_t = \epsilon_t \sim iid(0,\sigma^2).$$

$$\mathcal{M}_2 : \quad y_t = \frac{1}{\alpha}\boldsymbol{E}_t[y_{t+1}] + u_t, \quad u_t = \rho u_{t-1} + \epsilon_t \sim iid(0,\sigma^2).$$

- Solution:

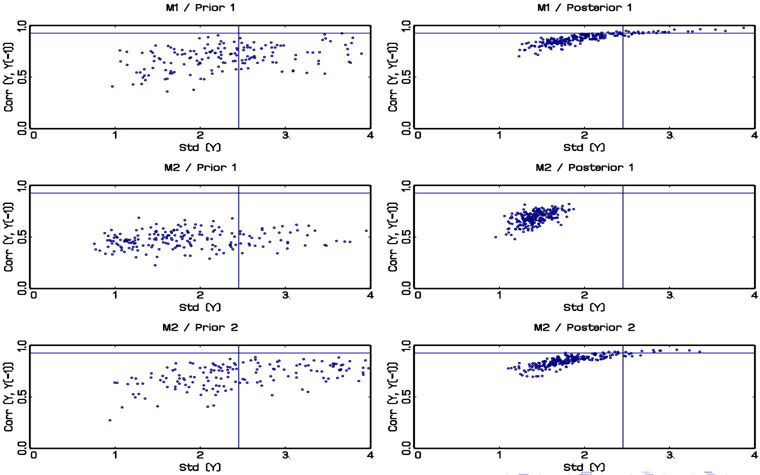
$$\mathcal{M}_1 : \quad y_t = \frac{1}{2}(\alpha - \sqrt{\alpha^2 - 4\rho\alpha})y_{t-1} + \frac{2\alpha}{\alpha + \sqrt{\alpha^2 - 4\rho\alpha}}\epsilon_t,$$

$$\mathcal{M}_2 : \quad y_t = \rho y_{t-1} + \frac{1}{1-\rho/\alpha}\epsilon_t$$

- Lubik and Schorfheide, Bayer and Farmer.

Priors and Model Comparisons in the Simple Example

1. Use same prior for $\mathcal{M}_1$ and $\mathcal{M}_2$
2. Choose prior for and $\mathcal{M}_2$ so that the same "a priori" implications for moments of the endogenous variables.



Priors and Model Comparisons in the Simple Example

1. Use **same prior** for $\mathcal{M}_1$ and $\mathcal{M}_2$

2. Choose prior for and $\mathcal{M}_2$ so that the **same** "a priori" **implications for moments** of the endogenous variables.

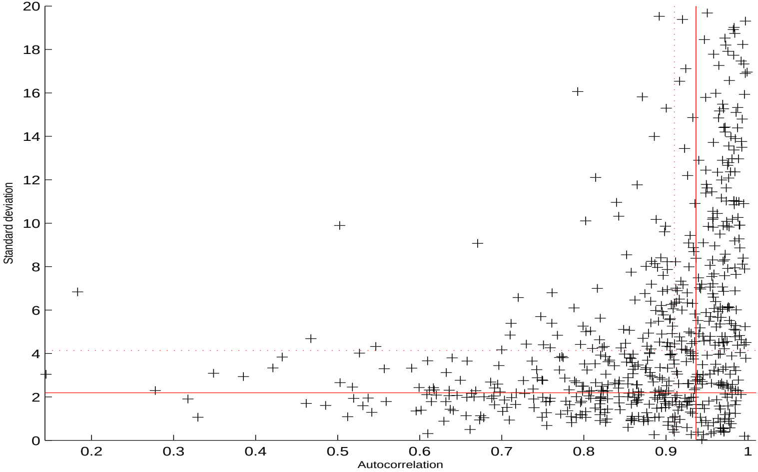
Specification	$\ln p(Y)$
Model $\mathcal{M}_1$, Prior 1	-105.93
Model $\mathcal{M}_2$, Prior 1	-123.53
Model $\mathcal{M}_2$, Prior 2	-105.70
Model $\mathcal{M}_1$, Prior 3	-108.93
Model $\mathcal{M}_2$, Prior 3	-108.24

Standard Practice in Bayesian DSGE Model Comparisons

1. Choose the priors from . . .

Standard Practice in Bayesian DSGE Model Comparisons

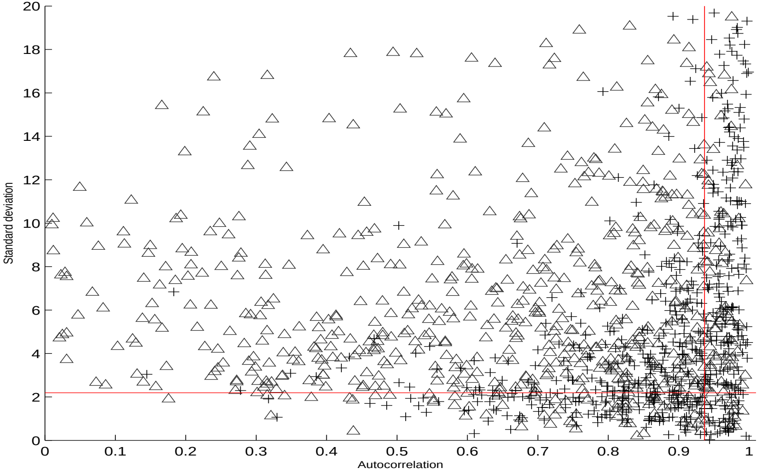
1. Choose the priors from . . . Smets and Wouters!
2. Use the same prior for all models considered.



Standard Practice in Bayesian DSGE Model Comparisons

1. Choose the priors from . . . Smets and Wouters!

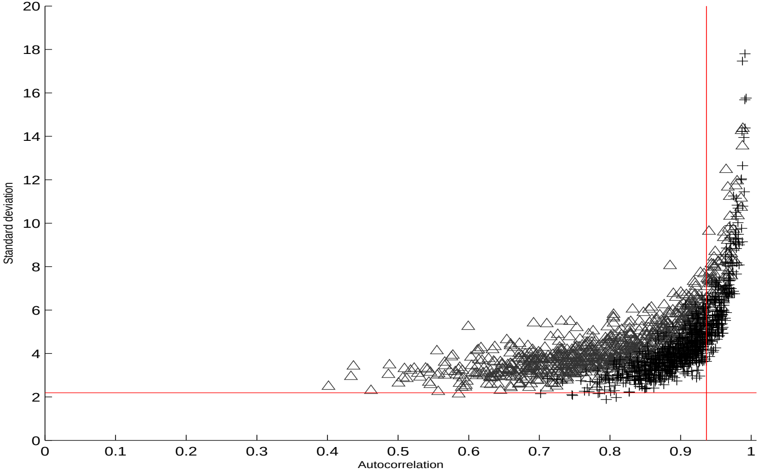
2. Use the same prior for all models considered.



Standard Practice in Bayesian DSGE Model Comparisons

1. Choose the priors from . . . Smets and Wouters!

2. Use the same prior for all models considered.



Adjusting Prior Distributions for Model Comparisons

1. Models $\mathcal{M}_i$, $i = 1, \ldots, J$ with parameter vectors $\theta^{(i)}$.

2. Split $\theta^{(i)}$ into $\theta^{(i)} = [\theta_1^{(i)} \; \theta_2^{(i)}]$ where θ_1 collects the "deep" parameters (prior distributions based on micro evidence) and θ_2 is a sub-vector of auxiliary parameters.

Adjusting Prior Distributions for Model Comparisons

1. Models $\mathcal{M}_i$, $i = 1, \ldots, J$ with parameter vectors $\theta^{(i)}$.

2. Split $\theta^{(i)}$ into $\theta^{(i)} = [\theta_1^{(i)} \; \theta_2^{(i)}]$ where θ_1 collects the "deep" parameters (prior distributions based on micro evidence) and θ_2 is a sub-vector of auxiliary parameters.

3. Pick a benchmark model (1) and a specific set of parameters $\underline{\theta}^{(1)}$ (say the prior mean), and compute the population covariance matrices of the endogenous variables: $\Gamma_{YY}(\underline{\theta}^{(1)})$ (shorthand notation $\Gamma_{YY}^{(1)}$) , $\Gamma_{XX}(\underline{\theta}^{(1)})$, etc.

Adjusting Prior Distributions for Model Comparisons . . .

4. Define the correction:

$$\mathcal{L}(\theta^{(i)}|\Gamma^{(1)}_{YY},\Gamma^{(1)}_{XY},\Gamma^{(1)}_{XX}) = |\Sigma_*(\theta^{(i)})|^{-(T^*+n+1)/2}$$
$$\times \exp\left\{-\frac{T^*}{2} tr\left[\Sigma_*(\theta^{(i)})^{-1}(\Gamma^{(1)}_{YY} - 2\Phi_*(\theta^{(i)})\Gamma^{(1)}_{XY} + \Phi'_*(\theta^{(i)})\Gamma^{(1)}_{XX}\Phi_*(\theta^{(i)})\right]\right\}$$

where $\Phi_*(\theta) = [\Gamma_{XX}]^{-1}\Gamma_{XY},\ \Sigma_*(\theta) = \Gamma_{YY} - \Gamma_{YX}[\Gamma_{XX}]^{-1}\Gamma_{XY}$.

Adjusting Prior Distributions for Model Comparisons . . .

4. Define the correction:

$$\mathcal{L}(\theta^{(i)}|\Gamma^{(1)}_{YY},\Gamma^{(1)}_{XY},\Gamma^{(1)}_{XX}) = |\Sigma_*(\theta^{(i)})|^{-(T^*+n+1)/2}$$
$$\times \exp\left\{-\frac{T^*}{2} tr\left[\Sigma_*(\theta^{(i)})^{-1}(\Gamma^{(1)}_{YY} - 2\Phi_*(\theta^{(i)})\Gamma^{(1)}_{XY} + \Phi'_*(\theta^{(i)})\Gamma^{(1)}_{XX}\Phi_*(\theta^{(i)})\right]\right\}$$

where $\Phi_*(\theta) = [\Gamma_{XX}]^{-1}\Gamma_{XY},\ \Sigma_*(\theta) = \Gamma_{YY} - \Gamma_{YX}[\Gamma_{XX}]^{-1}\Gamma_{XY}.$

5. Rather than the standard prior:

$$p(\theta_1,\theta_2) = \pi(\theta_1)\pi(\theta_2).$$

Use the **corrected** prior:

$$p_*(\theta_1,\theta_2) = \pi(\theta_1)\ \ c_1(.)\mathcal{L}(\underline{\theta}_1,\theta_2|\Gamma^{(1)})\pi(\theta_2).$$

where $c_1(.)$ guarantees the prior integrates to one.

. . . In Plain English

1. Generate artificial data from the benchmark model.
2. Estimate model (i) on this artificial data

. . . In Plain English

1. Generate artificial data from the benchmark model.

2. Estimate model (i) on this artificial data
   . . . fixing the “deep” parameters (θ_1) and letting only the auxiliary parameters (θ_2) vary.

... In Plain English

1. Generate artificial data from the benchmark model.
2. Estimate model (i) on this artificial data
   ... fixing the "deep" parameters (θ_1) and letting only the auxiliary parameters (θ_2) vary.

Outcome:

1. Make sure that for all models considered the auxiliary parameters are chosen so that the implications for second moments are as close as possible to the benchmark's.
2. Introduce correlation among auxiliary parameters.

DSGE Model

- Model is a variant of Altig, Christiano, Eichenbaum, and Linde (2002); Christiano, Eichenbaum, and Evans (2004), Smets and Wouters (2003).

- Continuum of households, they maximize:

$$\boldsymbol{E}_t \sum_{s=0}^{\infty} \beta^s [\log(C_{t+s} - hC_{t+s-1}) - \frac{\varphi_{t+s}}{1+\nu_l} L_{t+s}^{1+\nu_l} \ldots$$
$$\ldots + \frac{\chi}{1-\nu_m} \left(\frac{M_{t+s}}{Z_{t+s} P_{t+s}} \right)^{1-\nu_m}],$$

- Accumulate capital: $\bar{K}_t = (1-\delta)\bar{K}_{t-1} + \left(1 - S\left(\frac{I_t}{I_{t-1}}\right)\right) I_t,$

- Rent out "effective" capital $K_t = u_t \bar{K}_{t-1}$ and pay the utilization cost $a(u_t)\bar{K}_{t-1}$.

DSGE Model – continued

- Sticky wages: reset wages with probability $1 - \zeta_w$.
- Partial indexation: $W_{t+s} = \left(\Pi_{l=1}^{s} (\pi_* e^{\gamma})^{1-\iota_w} (\pi_{t+l-1} e^{\gamma})^{\iota_w}\right) \tilde{W}_t$.
- Continuum of intermediate goods producers, who use Cobb-Douglas technology:

$$Y_t(i) = K_t(i)^{\alpha} (Z_t L_t(i))^{1-\alpha}$$

with unit root in technology: $z_t = \log(Z_t / Z_{t-1})$ has mean γ.

DSGE Model – continued

- Sticky wages: reset wages with probability $1 - \zeta_w$.
- Partial indexation: $W_{t+s} = \left(\Pi_{l=1}^{s} (\pi_* e^{\gamma})^{1-\iota_w} (\pi_{t+l-1} e^{\gamma})^{\iota_w} \right) \tilde{W}_t$.
- Continuum of intermediate goods producers, who use Cobb-Douglas technology:

$$Y_t(i) = K_t(i)^{\alpha} (Z_t L_t(i))^{1-\alpha}$$

  with unit root in technology: $z_t = \log(Z_t / Z_{t-1})$ has mean γ.

- Sticky prices: reset prices with probability $1 - \zeta_p$ + Partial indexation (ι_p).
- $Y_t(i)$ packed into a composite good: $Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{1+\lambda_{f,t}}} di \right]^{1+\lambda_f}$.

DSGE Model – continued

- Government balances budget

$$P_t G_t + R_{t-1} B_{t-1} + M_{t-1} = T_t + M_t + B_t$$

where $G_t = (1 - 1/g_t) Y_t$.

- The central bank follows a nominal interest rate rule:

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*} \right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_*} \right)^{\psi_1} \left(\frac{Y_t}{Y_t^*} \right)^{\psi_2} \right]^{1-\rho_R} \sigma_R e^{\epsilon_{R,t}}$$

where Y_t^* is the stochastic steady state level of output.

- All shocks follow an AR(1) process (except $\epsilon_{R,t}$, which is iid).

Measurement equations

100 quarters of data ending Q1-2004.

- Output growth (log differences, quarter-to-quarter, in %):

$$100 \times (\ \ln Y_t - \ln Y_{t-1}\ ) \quad = 100 \times (\ \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t\ ) + 100\gamma$$

- Hours worked (log): $\ln L_t = \ 100 \times \hat{L}_t + \ln L^{adj}$

- Inflation (annualized, in %):

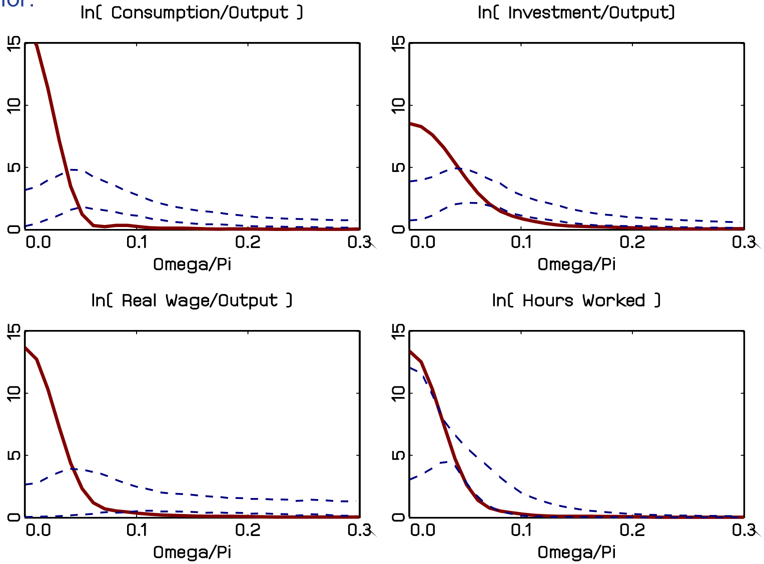
$$400 \times (\ \ln P_t - \ln P_{t-1}\ ) \quad = 400\hat{\pi}_t + 400 \ln \pi^*$$

- Nominal interest rate (annualized, in %):

$$400 \times (\ \ln R_t\ ) \quad = 4 \times 100\hat{R}_t + 400 * \ln R^*$$

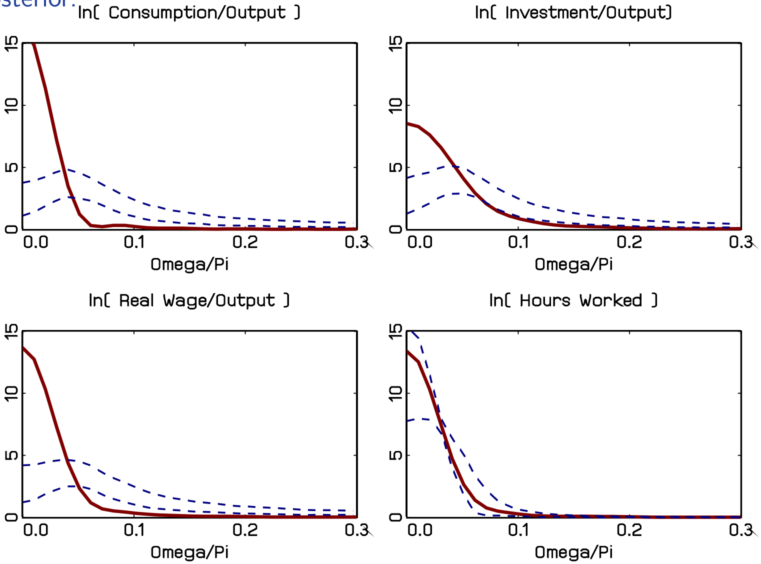
The “Big Ratios” and Hours Worked: Smoothed Periodograms for Model and Data

Prior:



The "Big Ratios" and Hours Worked: Smoothed Periodograms for Model and Data

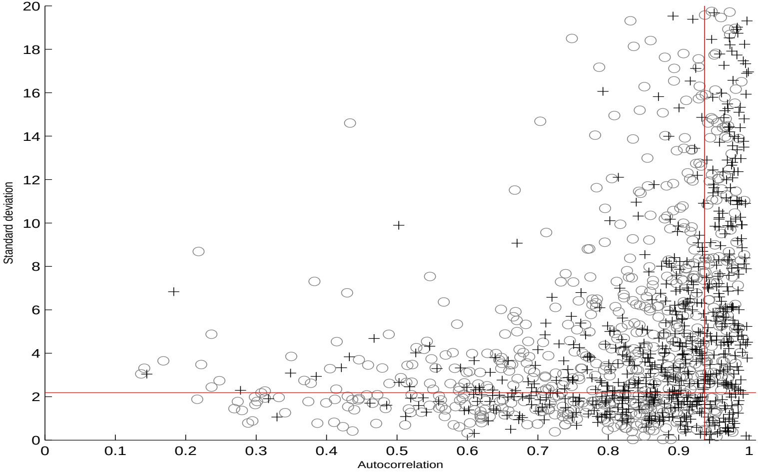
Posterior:



Some Examples

Baseline vs No Indexation

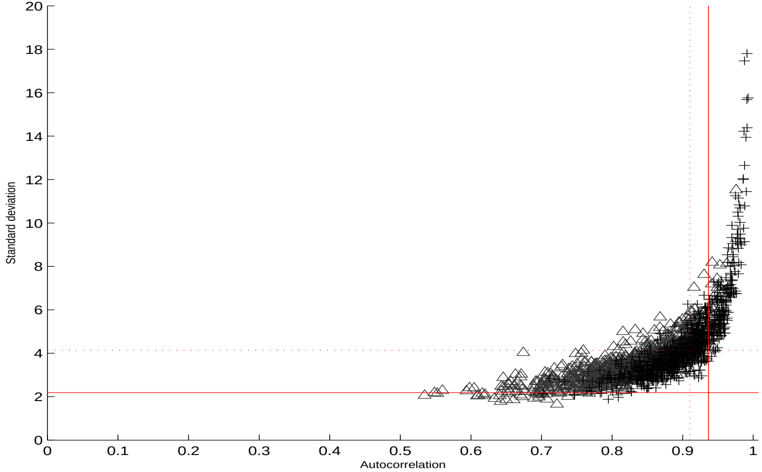
Before . . .



Some Examples

Baseline vs No Indexation

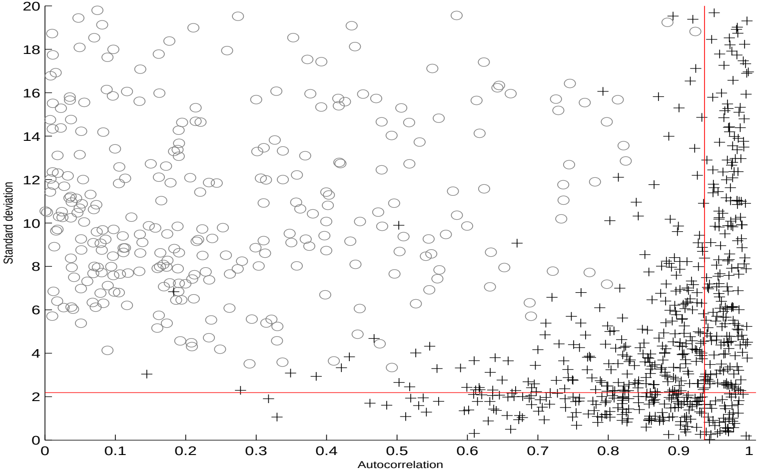
. . . and After!



Some Examples

Baseline vs Flexible Wages & Prices

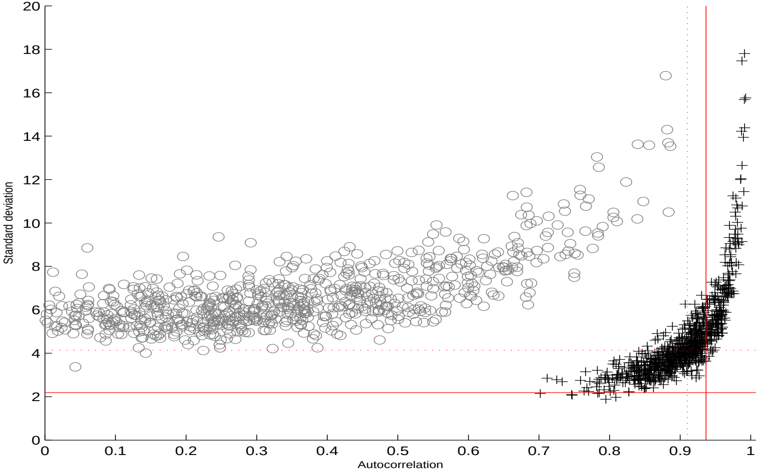
Before . . .



Some Examples

Baseline vs Flexible Wages & Prices

. . . and After!



Results: Log Marginal Data Densities Relative to Benchmark

Specification	$T^* = 0$	$T^* = 4$	$T^* = 25$
Benchmark	0.00	0.40	-3.93

Results: Log Marginal Data Densities Relative to Benchmark

Specification	$T^* = 0$	$T^* = 4$	$T^* = 25$
Benchmark	0.00	0.40	-3.93
Full Indexation	-10.41	-13.57	-25.07
No Indexation	2.01	3.80	-0.70

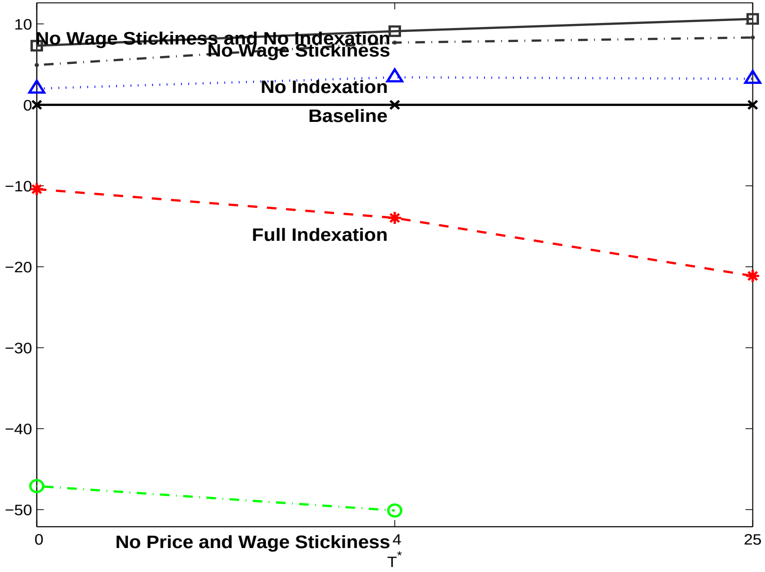
Results: Log Marginal Data Densities Relative to Benchmark

Specification	$T^* = 0$	$T^* = 4$	$T^* = 25$
Benchmark	0.00	0.40	-3.93
Full Indexation	-10.41	-13.57	-25.07
No Indexation	2.01	3.80	-0.70
No Price and Wage Stickiness	-47.09	-49.71	NaN

Results: Log Marginal Data Densities Relative to Benchmark

Specification	$T^* = 0$	$T^* = 4$	$T^* = 25$
Benchmark	0.00	0.40	-3.93
Full Indexation	-10.41	-13.57	-25.07
No Indexation	2.01	3.80	-0.70
No Price and Wage Stickiness	-47.09	-49.71	NaN
No Wage Stickiness	4.91	8.08	4.38
No Wage Stickiness and No Indexation	7.31	9.50	6.68

Log Marginal Data Densities "Fan"



Conclusion

- Priors matter

Conclusion

- Priors matter – using the same priors across different models may not be a good idea.

Conclusion

- Priors matter – using the same priors across different models may not be a good idea.
- Methodology for choosing “reasonable” priors for auxiliary parameters

Conclusion

- Priors matter – using the same priors across different models may not be a good idea.

- Methodology for choosing "reasonable" priors for auxiliary parameters – focusing on the implications for the volatilities and correlation of the observables.

    1. Introduce dependence among parameters.

    2. Levels the playing field for model comparisons – makes sure that the prior implications for the moments of the endogenous variables is the same across models.

Parameters – Baseline model

Parameter	Prior Mean	Prior Stdd	Post Mean	90% Lower Band	90% Upper Ban
ζ_p	0.600	0.200	0.684	0.581	0.786
ι_p	0.500	0.280	0.055	0.000	0.125
s'	4.000	1.500	8.790	5.753	11.814
h	0.700	0.050	0.759	0.678	0.840
$a'\,'$	0.200	0.100	0.175	0.038	0.312
ζ_w	0.600	0.200	0.124	0.023	0.219
ι_w	0.500	0.280	0.464	0.009	0.871
ψ_1	1.500	0.400	2.037	1.624	2.414
ψ_2	0.200	0.100	0.075	0.034	0.117
ρ_r	0.500	0.200	0.690	0.632	0.753
ρ_z	0.400	0.250	0.532	0.333	0.709
ρ_ϕ	0.750	0.250	0.978	0.952	1.000
ρ_g	0.750	0.250	0.915	0.856	0.983
σ_z	0.500	4.000	0.865	0.739	0.989
σ_ϕ	4.500	4.000	2.986	2.187	3.798
σ_g	0.750	4.000	0.625	0.522	0.731
σ_r	0.200	4.000	0.288	0.251	0.325

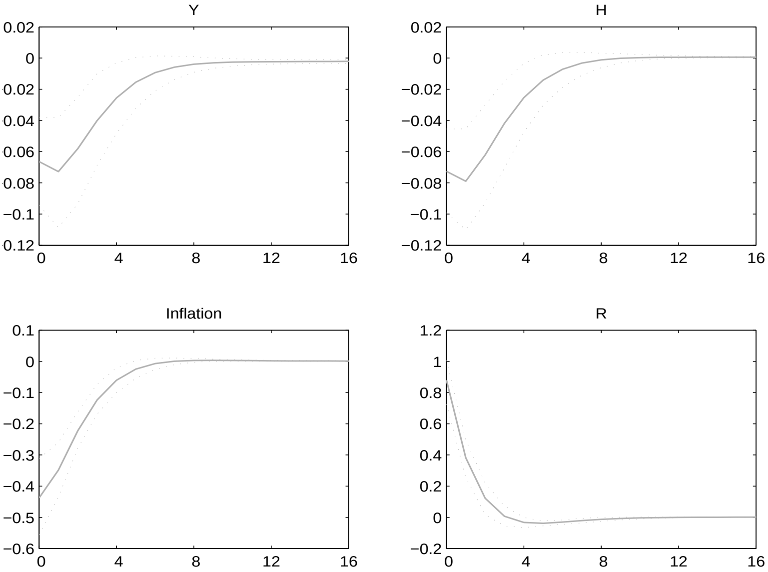
Parameters – Baseline model w/ correction

Parameter	Prior Mean	Prior Stdd	Post Mean	90% Lower Band	90% Upper Ba
ζ_p	0.600	0.200	0.736	0.663	0.810
ι_p	0.500	0.280	0.050	0.000	0.114
s'	4.000	1.500	8.351	5.361	11.373
h	0.700	0.050	0.740	0.659	0.823
$a' \, '$	0.200	0.100	0.130	0.021	0.236
ζ_w	0.600	0.200	0.144	0.023	0.261
ι_w	0.500	0.280	0.474	0.005	0.877
ψ_1	1.500	0.400	1.931	1.545	2.304
ψ_2	0.200	0.100	0.086	0.042	0.130
ρ_r	0.500	0.200	0.717	0.661	0.771
ρ_z	0.400	0.250	0.266	0.065	0.465
ρ_ϕ	0.750	0.250	0.951	0.905	1.000
ρ_g	0.750	0.250	0.894	0.841	0.947
σ_z	0.500	4.000	0.773	0.681	0.866
σ_ϕ	4.500	4.000	3.167	2.418	3.923
σ_g	0.750	4.000	0.803	0.699	0.912
σ_r	0.200	4.000	0.277	0.245	0.308

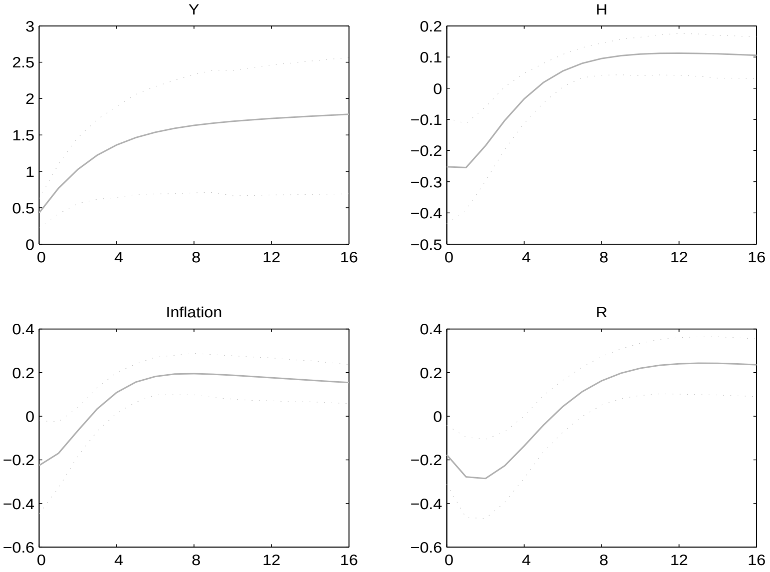
Parameters – No Wage Rigidity & Ind.

Parameter	Prior Mean	Prior Stdd	Post Mean	90% Lower Band	90% Upper Ban
ζ_p	0.600	0.200	0.767	0.716	0.817
s'	4.000	1.500	9.062	6.070	12.039
h	0.700	0.050	0.775	0.704	0.851
$a' \, '$	0.200	0.100	0.225	0.068	0.369
ψ_1	1.500	0.400	2.039	1.665	2.390
ψ_2	0.200	0.100	0.074	0.034	0.113
ρ_r	0.500	0.200	0.680	0.621	0.740
ρ_z	0.400	0.250	0.456	0.313	0.596
ρ_ϕ	0.750	0.250	0.979	0.957	1.000
ρ_g	0.750	0.250	0.933	0.878	1.000
σ_z	0.500	4.000	0.842	0.736	0.949
σ_ϕ	4.500	4.000	2.755	2.041	3.411
σ_g	0.750	4.000	0.641	0.540	0.734
σ_r	0.200	4.000	0.295	0.256	0.332

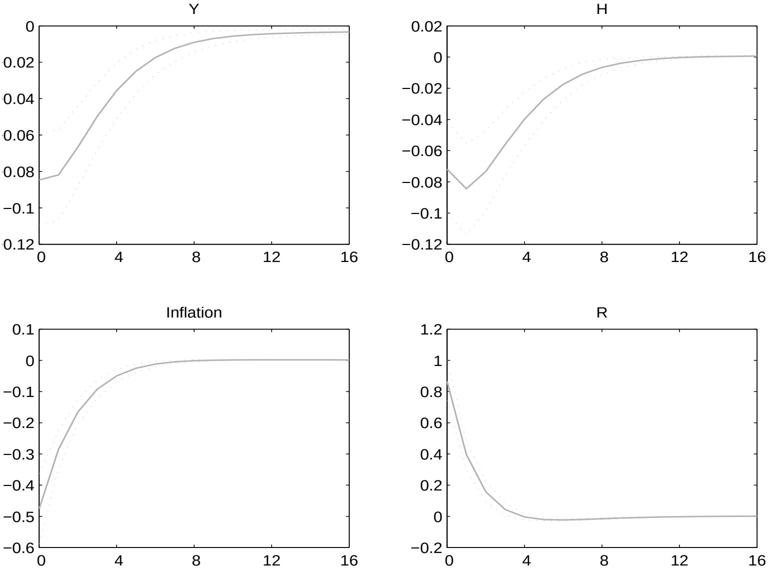
IRFs Money – Baseline



IRFs Tech – Baseline



IRFs Money – No Wage Rigidity & Ind.



IRFs Money – No Wage Rigidity & Ind.