Back to Square One: Identification Issues in DSGE Models

Fabio Canova
ICREA-UPF, CREI, AMeN and CEPR

Luca Sala
IGIER, Universitá Bocconi

Bank of Finland, June 2006

- DSGE models are now regularly and successfully estimated

- inference on a large number of parameters is conducted

$\rightarrow$ DSGE models have quickly become the benchmark for:

- Understanding business cycles

- Policy analysis

How are DSGE estimated?

Full Information

- Maximum Likelihood

- Bayesian Maximum Likelihood (weighted average of prior and likelihood)

Limited Information

- Indirect Inference: minimum distance

$\rightarrow$ matching impulse responses

Maximum Likelihood and Indirect Inference

Get the unique stationary rational expectations equilibrium $\longrightarrow$ (restricted VAR):

$$A(\theta_0)Y_t = B(\theta_0)Y_{t-1} + u_t$$

or:

$$Y_t = C(\theta_0)Y_{t-1} + D(\theta_0)u_t$$

ML: $\hat{\theta} = argmax \; L(\theta, Y)$

Bayesian ML: $\hat{\theta} = argmax \; L(\theta, Y)\pi(\theta)$

Matching impulse responses (conditional on shock j):

1) Estimate a VAR, identify shock j, compute impulse responses

$$\underbrace{\widehat{Y}_t^{D,j}}_{horN\times 1} = vec\ \underbrace{\widehat{W}(L)\widehat{u}_t^j}_{hor\times N} \qquad \rightarrow \text{ you need instruments!!}$$

2) Compute model impulse responses to shock j:

$$\underbrace{Y_t^{M,j}(\theta)}_{horN\times 1} = vec\ \underbrace{P(\theta)(L)u_t^j}_{hor\times N}$$

3) $\widehat{\theta} = argmin\ D(\theta)$

where: $D(\theta) = ||\widehat{Y}_t^{D,j} - Y_t^{M,j}(\theta)||_{\widehat{\Omega}}$ ($\widehat{\Omega}$: inverse of std.err. of $\widehat{Y}_t^{D,j}$)

Question:

Under what conditions can we recover structural parameters?

Identifiability

Mapping from the loss function to the structural parameters

- Likelihood (Rothenberg, 1971):

θ_0 is *identifiable* if $\nexists$ any θ_1 such that: $L(y, \theta_0) = L(y, \theta_1)$ for all y

- Distance:

- $D(\theta)$ has a *unique minimum* 0 at $\theta = \theta_0$
- Hessian is *positive definite* and *full rank* at $\theta = \theta_0$

<u>Additional issue</u>

"Enough" curvature in $D(\theta)$ and $L(y, \theta)$ ($\Delta\theta \rightarrow \Delta D(\theta)$ sufficient)

$\longrightarrow$ analog to weak instruments in IV and GMM

Consequences of under and weak identification

- Linear IV setup (Choi-Phillips (92), Staiger-Stock (97))

1. parameter estimates inconsistent
2. asymptotic distributions non-standard
3. standard tests incorrect

- Same in GMM setups (Stock and Wright (2000))

$\longrightarrow$ Similar problems in DSGE?

Additonal practical problem in DSGE:

- Remain stuck in local minima if algorithm is poor

Different loss functions may have different ”identification power”

In DSGE, the shape of the likelihood $L(.)$ and distance $D(.)$ is too complicated to be worked out analitically

$\downarrow$

Identifiability is far from clear

We study the shape of the loss functions

Results

- Some *model-loss function* combinations display some identification problems

- Warning on the interpretation of the estimated parameters

- Warning on the structural interpretation of the model

- Proposals for applied researchers

Example 1

$$
\begin{aligned}
y_t &= a_1 E_t y_{t+1} + a_2(i_t - E_t \pi_{t+1}) + v_{1t} && (1)\\
\pi_t &= a_3 E_t \pi_{t+1} + a_4 y_t + v_{2t} && (2)\\
i_t &= a_5 E_t \pi_{t+1} + v_{3t} && (3)
\end{aligned}
$$

Unique RE solution (log-lin from steady state):

$$
\begin{bmatrix} \widehat{y}_t \\ \widehat{\pi}_t \\ \widehat{i}_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & a_2 \\ a_4 & 1 & a_2 a_4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{bmatrix}
$$

- Some parameters disappear from the solution

- Different shocks identify different parameters

- ML and distance have different identification properties.

Example 2: a Neo-Keynesian model

$$
\begin{aligned}
y_t &= \frac{h}{1+h} y_{t-1} + \frac{1}{1+h} E_t y_{t+1} - \frac{1}{\phi}(i_t - E_t \pi_{t+1}) + v_{1t} \\
\pi_t &= \frac{\omega}{1+\omega\beta} \pi_{t-1} + \frac{\beta}{1+\omega\beta} \pi_{t+1} + \frac{(\phi+\nu)(1-\zeta\beta)(1-\zeta)}{(1+\omega\beta)\zeta} y_t + v_{2t} \\
i_t &= \lambda_r i_{t-1} + (1-\lambda_r)(\phi_\pi \pi_{t-1} + \lambda_y y_{t-1}) + v_{3t}
\end{aligned}
$$

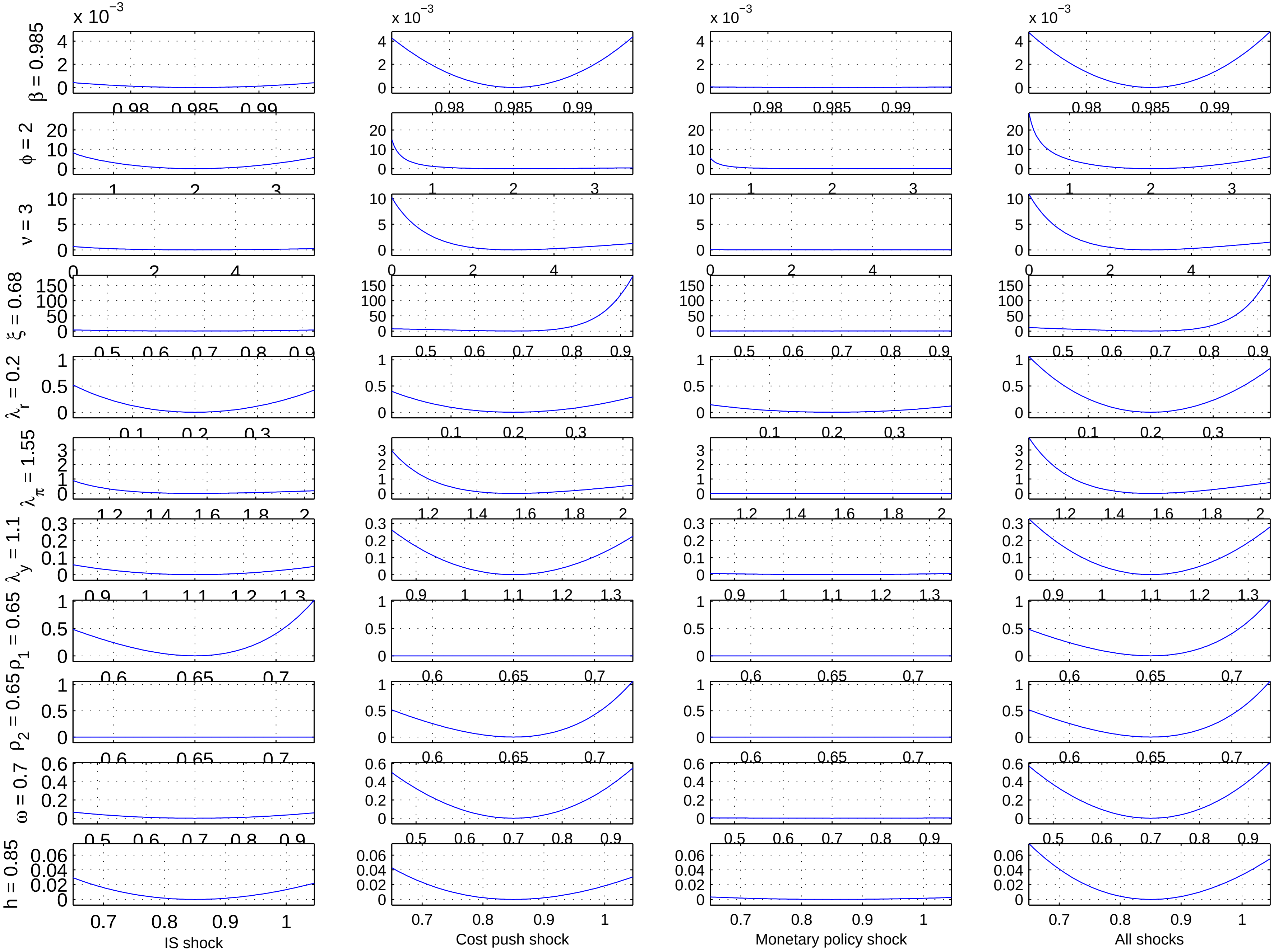
h: degree of habit persistence	(.85)
ν: inverse elasticity of labor supply	(3)
ϕ: relative risk aversion	(2)
β: discount factor	(.985)
ω: degree of price indexation	(.25)
ζ: degree of price stickiness	(.68)

$\lambda_r, \lambda_\pi, \lambda_y$: policy parameters (.2, 1.55, 1.1)
v_{1t}: AR(ρ_1) (.65)
v_{2t}: AR(ρ_2) (.65)
v_{3t}: i.i.d.

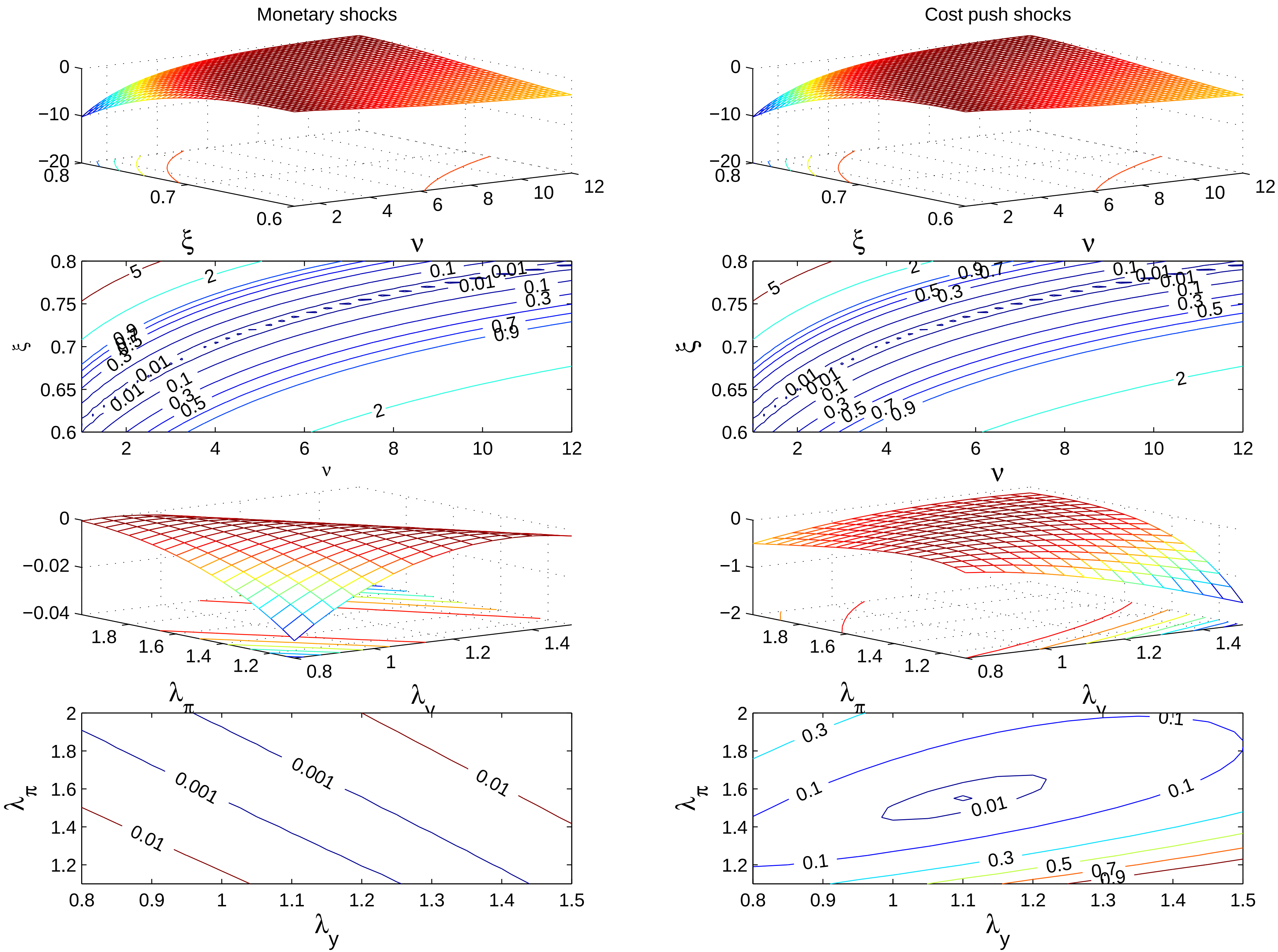
Compute IRFs for: $[\pi_t, i_t, y_t]$ to v_{it}: $Y_t^M(\theta_0)$

Loss: $D(\theta) = ||\hat{Y}_t^D - Y_t^M(\theta)||_{\hat{\Omega}} = ||Y_t^M(\theta_0) - Y_t^M(\theta)||$

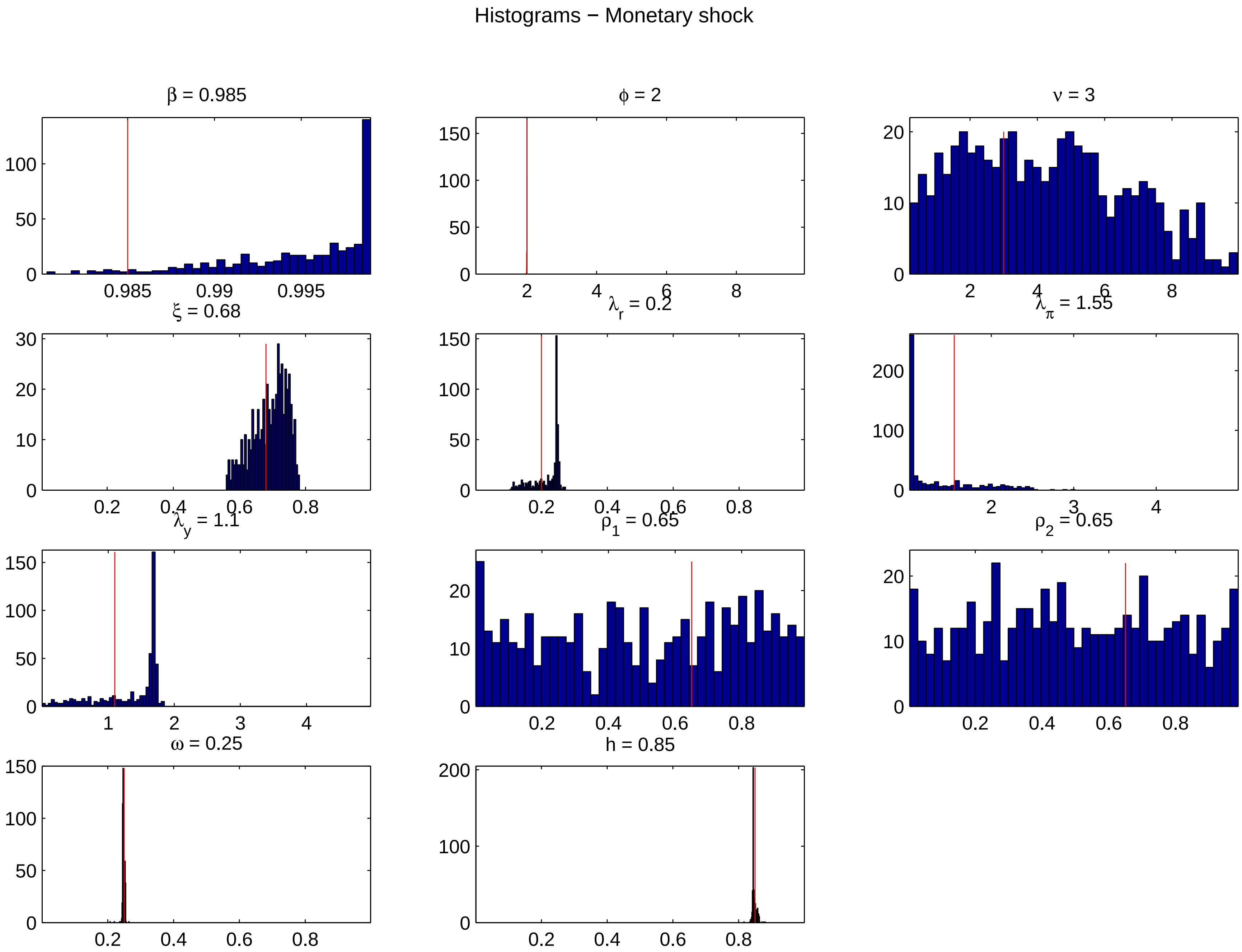
Plot the population loss function as θ varies



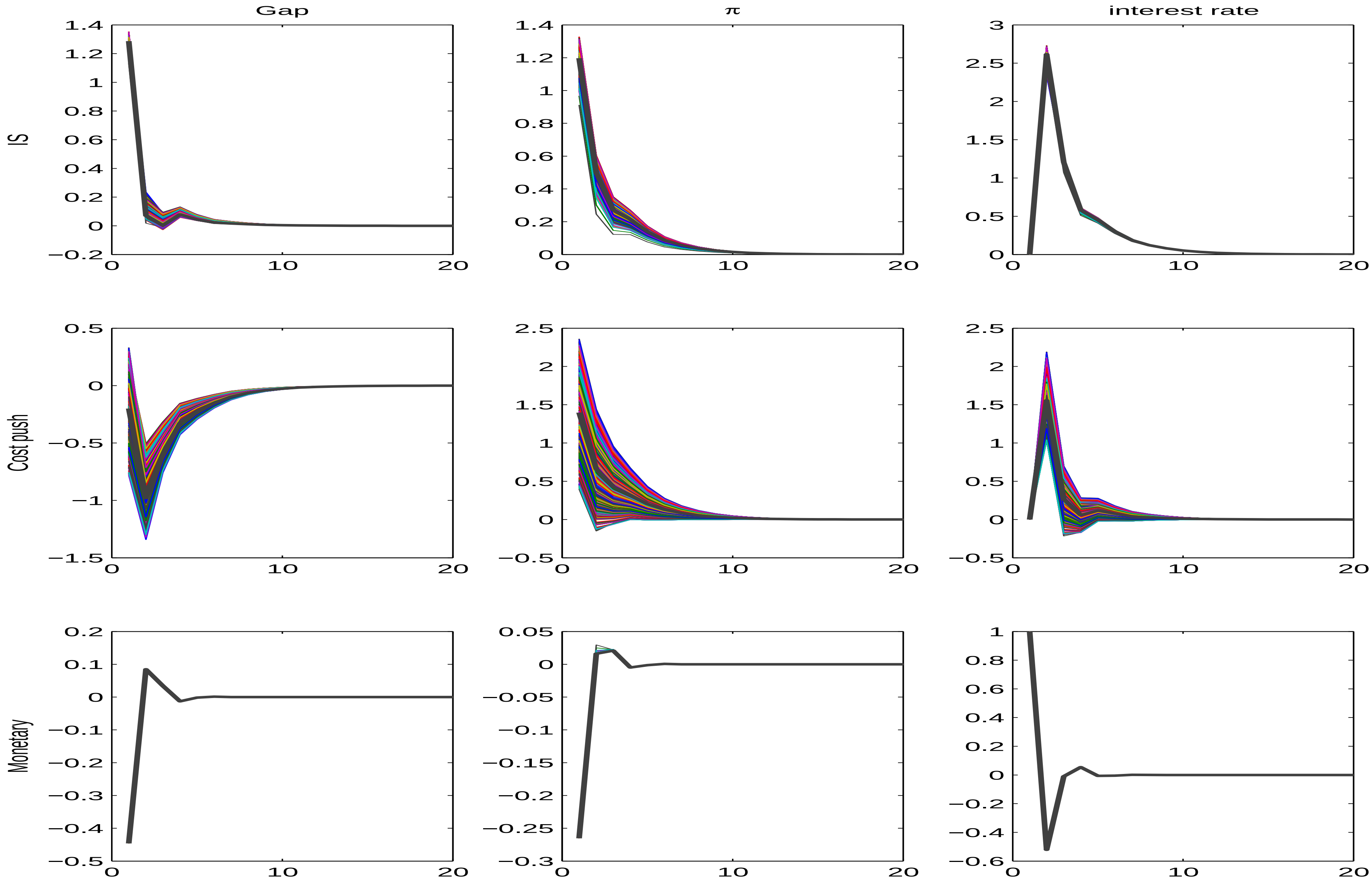
β = 0.985
φ = 2
ν = 3
ξ = 0.68
λr = 0.2
λπ = 1.55
λy = 1.1
ρ1 = 0.65
ρ2 = 0.65
ω = 0.7
h = 0.85
x 10−3
IS shock
Cost push shock
Monetary policy shock
All shocks



Monetary shocks
Cost push shocks
ξ
ν
λπ
λy



Histograms − Monetary shock
β = 0.985
φ = 2
ν = 3
ξ = 0.68
λr = 0.2
λπ = 1.55
λy = 1.1
ρ1 = 0.65
ρ2 = 0.65
ω = 0.25
h = 0.85



Gap
π
interest rate
IS
Cost push
Monetary

Small sample properties

- simulate data 500 times, estimate a VAR
- identify (correctly) the monetary policy shock
- compute weighting matrix $\widehat{\Omega}$ by bootstrap
- recover θ by minimizing $D(\theta)$, using $\widehat{\Omega}$

NK model. Matching monetary policy shocks

	True	T = 120	T = 200	T=1000	T=1000 wrong
β	.985	.98 (.007)	.98 (.006)	.98 (.007)	.999 (.008)
ϕ	2	1.49 (2.878)	1.504 (1.906)	1.757 (.823)	10 (.420)
ν	3	4.184 (1.963)	4.269 (1.763)	4.517 (1.634)	1.421 (2.33)
ζ	.68	.644 (.156)	.641 (.112)	.621 (.071)	.998(.072)
λ_r	.2	.552 (.272)	.481 (.266)	.352 (.253)	.417 (.099)
λ_π	1.55	1.058 (1.527)	1.107 (1.309)	1.345 (1.186)	3.607 (1.281)
λ_y	1.1	4.304 (2.111)	2.924 (2.126)	1.498 (2.088)	2.59 (1.442)
ρ_1	.65	.5 (.209)	.5 (.212)	.5 (.167)	.5 (.188)
ρ_2	.65	.5 (.208)	.5 (.213)	.5 (.188)	.5 (.193)
ω	.25	1 (.360)	1 (.35)	1 (.306)	0 (.384)
h	.85	1 (.379)	1 (.321)	1 (.233)	0 (.166)

- Large biases
- Std. err. do not become more precise as T increases
- Consistency?

Example 3: Calibrating some parameters in the RBC model

$$max \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\phi}}{1-\phi}$$

$$c_t + k_t = k_{t-1}^{\eta} z_t + (1-\delta)k_{t-1}$$

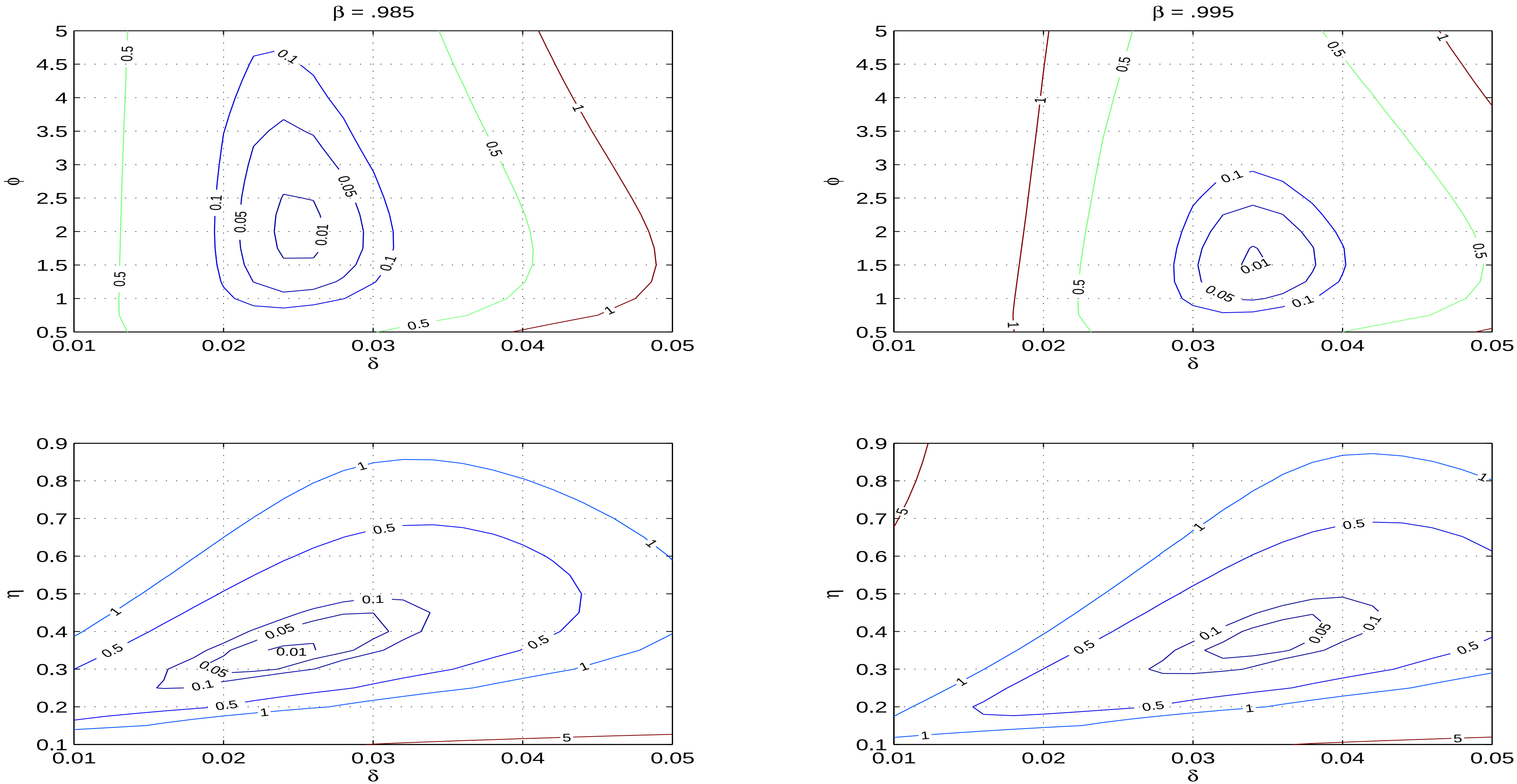
$$z_t = (1-\rho)z^{ss} + \rho z_{t-1} + e_t$$

$$\beta = .985, \phi = 2, \rho = .95, \eta = .36, \delta = .025, z^{ss} = 1$$

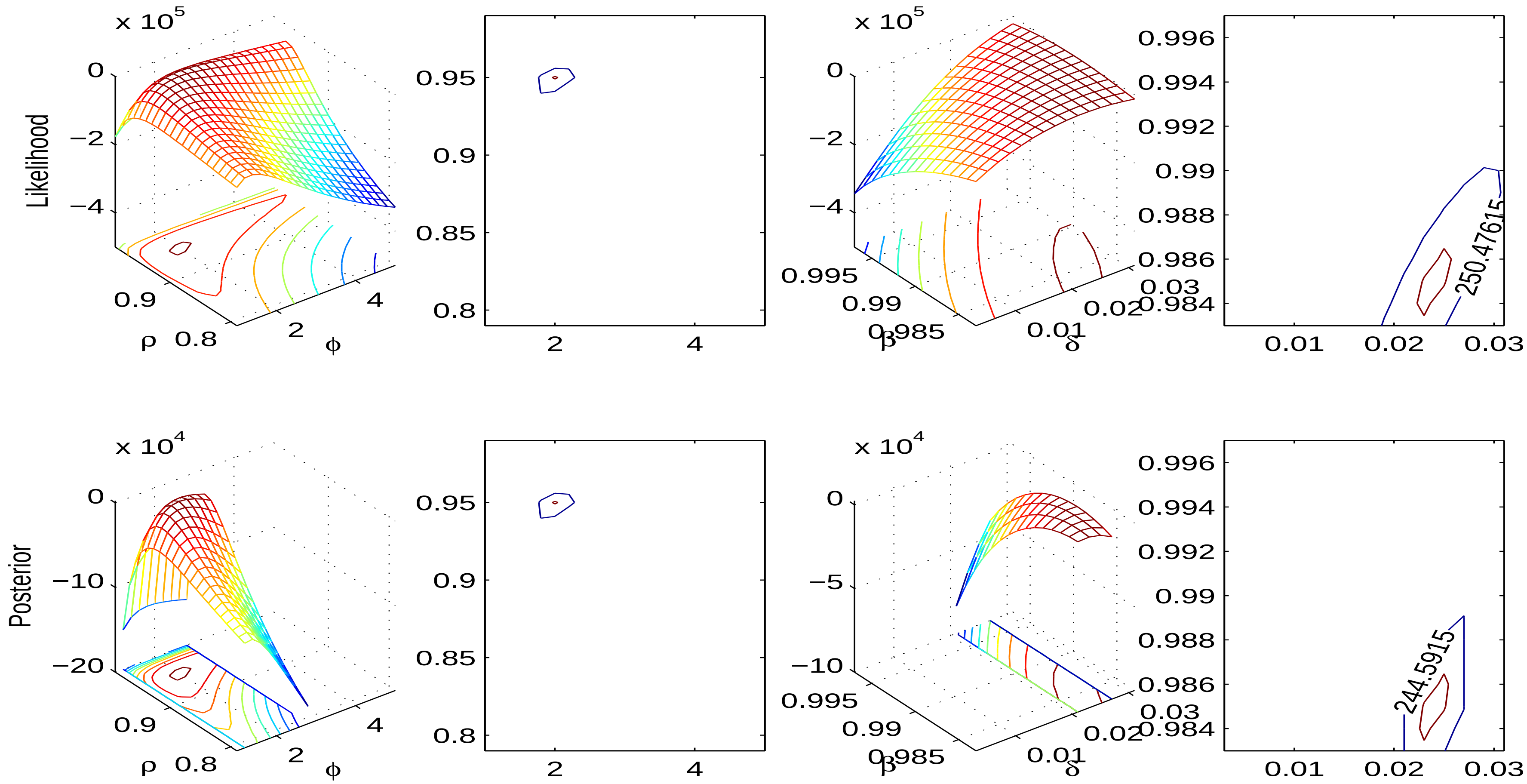
Compute IRFs for: $[k_t, c_t, y_t, z_t, r_t]$ to e_t: $Y_t^M(\theta_0)$

β is calibrated at .995

Calibrating β: shifts in the distribution



Likelihood and priors



Bayesian identification (Poirier, 1998)

• If parameter space is *variation free*: no updating of unidentified parameters

• If parameter space is not *variation free (or if prior correlated)*, as in DSGE: updating also of unidentified parameters

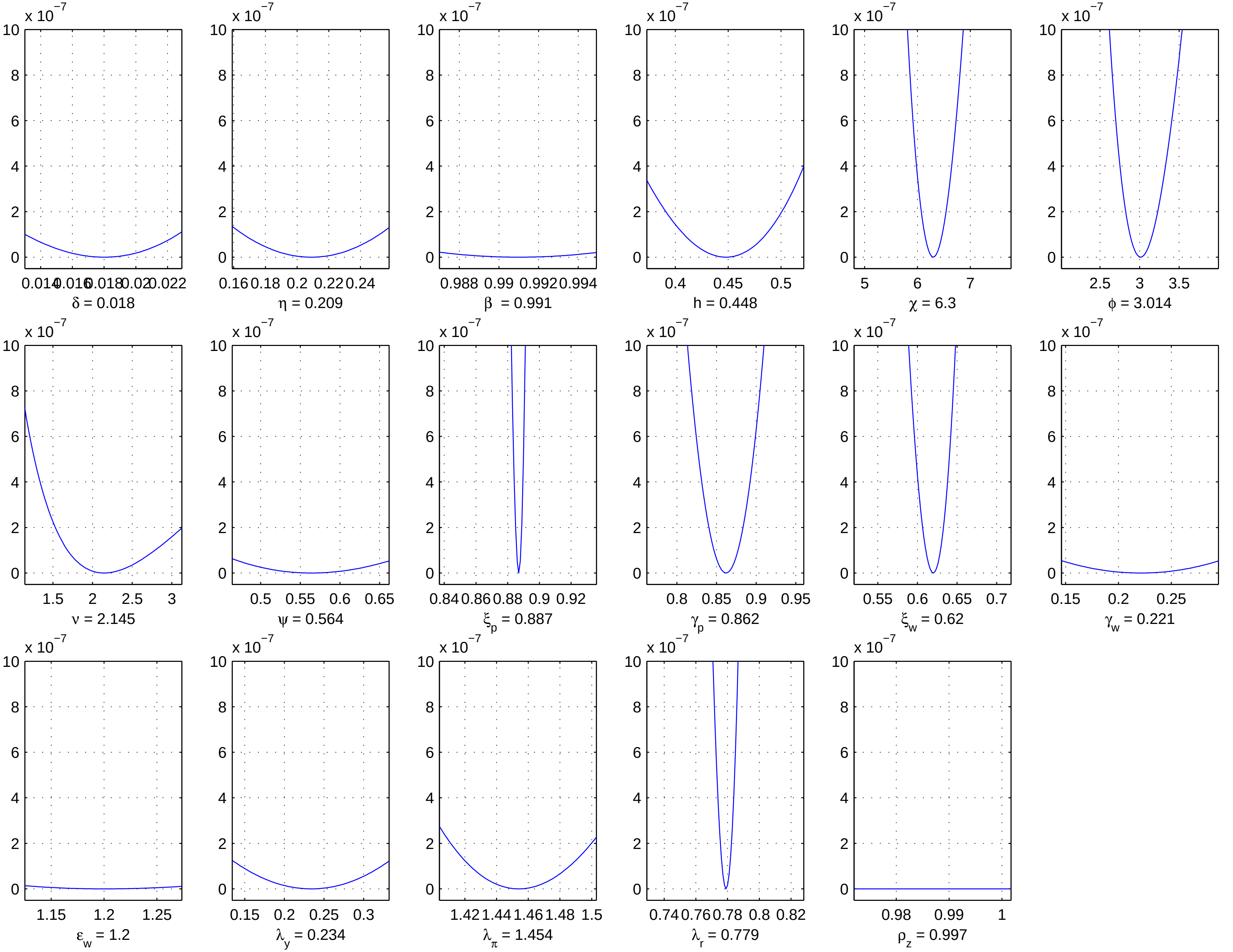
$\longrightarrow$ Robustness w.r.t. priors

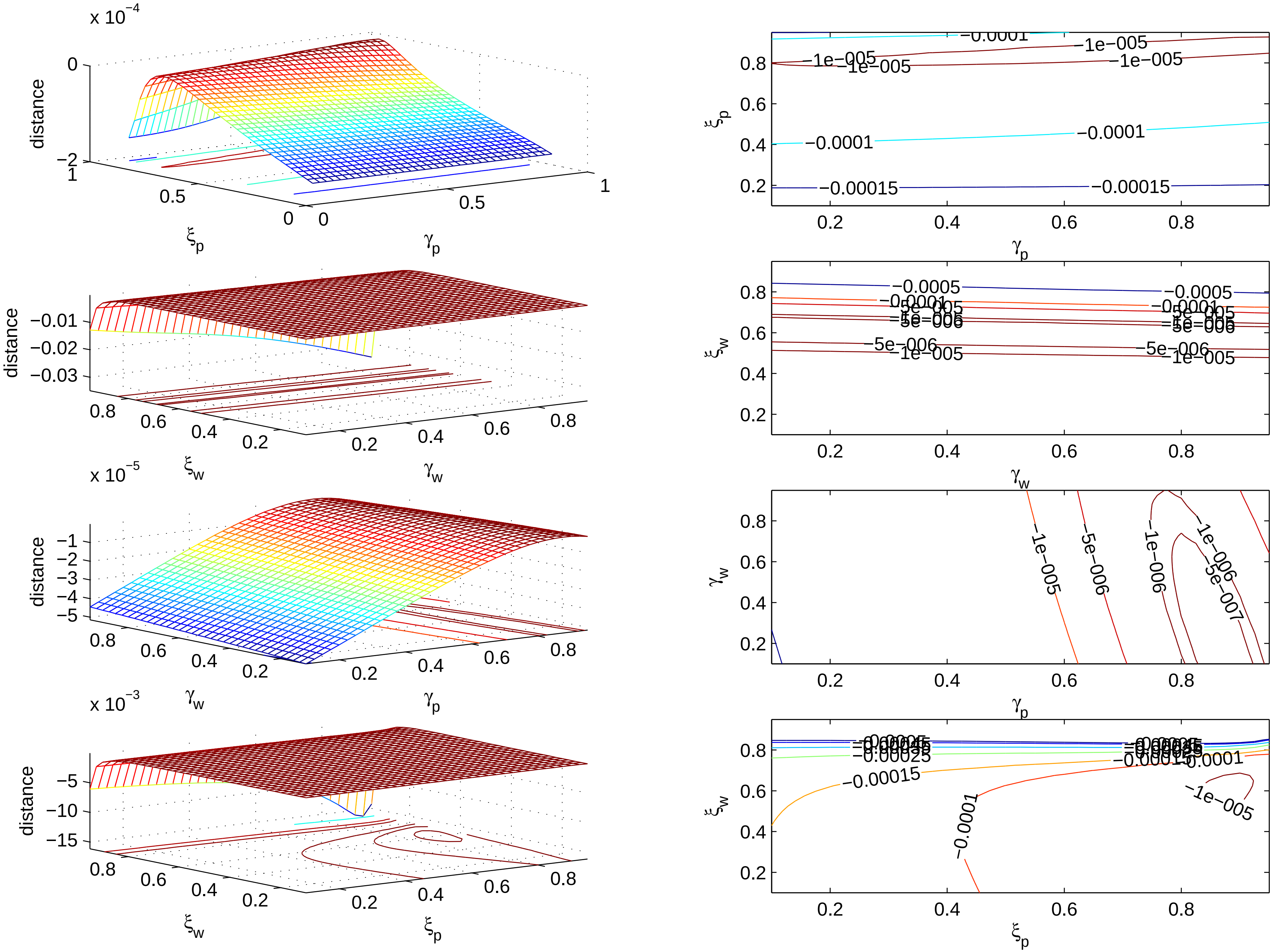
Example 4: a state-of-the-art DSGE model

$$
\begin{aligned}
0 &= -k_{t+1} + (1-\delta)k_t + \delta x_t \\
0 &= -u_t + \psi r_t \\
0 &= \frac{\eta\delta}{\bar{r}} x_t + (1 - \frac{\eta\delta}{\bar{r}}) c_t - \eta k_t - (1-\eta) N_t - \eta u_t - e z_t \\
0 &= -R_t + \phi_r R_{t-1} + (1-\phi_r)(\phi_\pi \pi_t + \phi_y y_t) + e r_t \\
0 &= -y_t + \eta k_t + (1-\eta) N_t + \eta u_t + e z_t \\
0 &= -N_t + k_t - w_t + (1+\psi) r_t \\
0 &= E_t[\frac{h}{1+h} c_{t+1} - c_t + \frac{h}{1+h} c_{t-1} - \frac{1-h}{(1+h)\varphi}(R_t - \pi_{t+1})] \\
0 &= E_t[\frac{\beta}{1+\beta} x_{t+1} - x_t + \frac{1}{1+\beta} x_{t-1} + \frac{\chi^{-1}}{1+\beta} q_t + \frac{\beta}{1+\beta} e x_{t+1} - \frac{1}{1+\beta} e x_t] \\
0 &= E_t[\pi_{t+1} - R_t - q_t + \beta(1-\delta) q_{t+1} + \beta \bar{r} r_{t+1}] \\
0 &= E_t[\frac{\beta}{1+\beta\gamma_p} \pi_{t+1} - \pi_t + \frac{\gamma_p}{1+\beta\gamma_p} \pi_{t-1} + T_p(\eta r_t + (1-\eta) w_t - e z_t + e p_t)] \\
0 &= E_t[\frac{\beta}{1+\beta\gamma_p} w_{t+1} - w_t + \frac{1}{1+\beta} w_{t-1} + \frac{\beta}{1+\beta} \pi_{t+1} - \\
&\quad \frac{1+\beta\gamma_w}{1+\beta} \pi_t + \frac{\gamma_w}{1+\beta\gamma_w}{}_{t-1} (w_t - \sigma N_t - \frac{\varphi}{1-h}(c_t - h c_{t-1}) - e w_t)]
\end{aligned}
$$

Parameters' values (from Dedola and Neri)

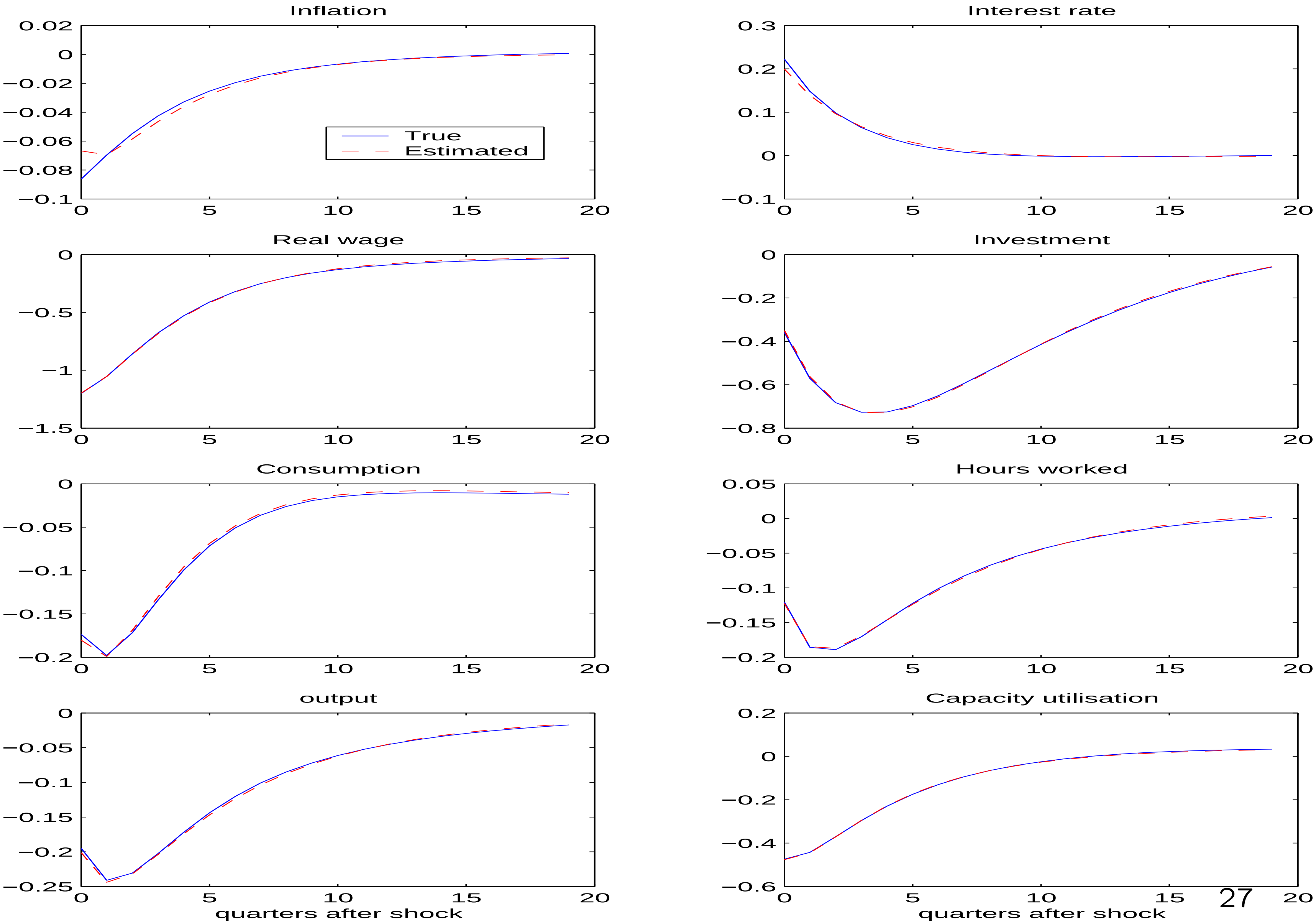
δ	depreciation rate	.0182	λ_w	wage markup	1.2
ψ	parameter	.564	$\bar{\pi}$	steady state π	1.016
η	share of capital	.209	h	habit persistence	.448
φ	risk aversion coefficient	3.014	σ_l	inverse elasticity of labor supply	2.145
β	discount factor	.991	χ^{-1}	investment's elasticity to Tobin's q	.15
ζ_p	price stickiness	.887	ζ_w	wage stickiness	.62
γ_p	price indexation	.862	γ_w	wage indexation	.221
ϕ_y	response to y	.234	ϕ_π	response to π	1.454
ϕ_r	int. rate smoothing	.779			
T_p	$\equiv \frac{(1-\beta\zeta_p)(1-\zeta_p)}{(1+\beta\gamma_p)\zeta_p}$				
T_w	$\equiv \frac{(1-\beta\zeta_w)(1-\zeta_w)}{(1+\beta)(1+(1+\lambda_w)\sigma_l\lambda_w^{-1})\zeta_w}$				


x 10^−7
δ = 0.018
η = 0.209
β = 0.991
h = 0.448
χ = 6.3
ϕ = 3.014
ν = 2.145
ψ = 0.564
ξp = 0.887
γp = 0.862
ξw = 0.62
γw = 0.221
εw = 1.2
λy = 0.234
λπ = 1.454
λr = 0.779
ρz = 0.997



x 10^−4
distance
ξp
γp
−0.0001
−1e−005
−0.00015
distance
ξw
γw
−0.0005
−5e−006
−1e−005
x 10^−5
distance
γw
γp
−1e−005
−5e−006
−1e−006
−5e−007
x 10^−3
distance
ξw
ξp
−0.00015
−0.0001
−0.00025
−1e−005

	ζ_p (p stick)	γ_p (p index)	ζ_w (w stick)	γ_w (w index)	Obj.Fun.
Baseline	**0.887**	**0.862**	**0.62**	**0.221**	
x0 = lb + 2std	0.8924	0.7768	0.6095	0.1005	3.75E-07
x0 = ub - 2std	0.9044	0.7701	0.6301	0	8.72E-07
Case 1	**0**	**0.862**	**0.62**	**0.221**	
x0 = lb + 2std	0.1015	0.0853	0.6065	0.1791	4.84E-08
x0 = ub - 2std	0.0922	0.0749	0.618	0.215	3.05E-08
Case 2	**0**	**0**	**0.62**	**0.221**	
x0 = lb + 2std	0.0838	0.1193	0.6044	0.1683	4.38E-08
x0 = ub - 2std	0.0789	0.0971	0.6114	0.1835	2.61E-08
Case 3	**0**	**0.862**	**0.62**	**0**	
x0 = lb + 2std	0.4649	0	0.7443	0.4668	2.10E-06
x0 = ub - 2std	0.6463	0.2673	0.8222	0.3811	5.56E-06
Case 4	**0.887**	**0**	**0.62**	**0.8**	
x0 = lb + 2std	0.9076	0.2268	0.6415	0.154	3.51E-07
x0 = ub - 2std	0.9263	0.3133	0.6294	0.4252	4.13E-07
Case 5	**0.887**	**0**	**0**	**0.221**	
x0 = lb + 2std	0.8994	0.234	0	0	3.06E-07
x0 = ub - 2std	0.9343	0.5409	0.0042	0	9.64E-07
Case 6	**0.887**	**0**	**0**	**0.221**	
x0 = lb + 2std	0.8919	0.0411	0.0003	0	4.26E-07
x0 = ub - 2std	0.8839	0.0499	0.0189	0	2.46E-06
Case 7	**0.887**	**0**	**0**	**0.221**	
x0 = lb + 2std	0.9052	0.2805	0	0.25	2.41E-07
x0 = ub - 2std	0.8985	0.194	0.001	0.25	2.07E-07


Inflation
True
Estimated
Interest rate
Real wage
Investment
Consumption
Hours worked
output
Capacity utilisation
quarters after shock
quarters after shock

Conclusions (1)

- Liu (1960), Sims (1980):

- Traditional models hopelessly under-identified.

- Identification often achieved not because we have sufficient information but because we want it to be so.

- Proceed with reduced form models

Conclusions (2)

A destructive approach...

- Models are identified not because loss f. + data is informative, but because we make it informative (via calibration or tight priors)

- Estimation = sophisticated calibration

Conclusions (3)

... and a more constructive one:

- Graphical analysis (univariate-bivariate)
- Study of the model and of its empirical properties
  - Reduced form (weak instrument test: Staiger-Stock, 97, Stock-Yogo, 01)
  - Mapping from reduced form to structural parameters
- Test for the rank of Hessian (Cragg-Donald, 1996, Anderson, 1951)
- Sensitivity to priors
- Be virtuous!!! Resist temptation ..... and (eventually) you will go to paradise......