



The European Commission's science and knowledge service

Joint Research Centre

Monte Carlo filtering and regional sensitivity analysis (RSA)

M. Ratto
DYNARE Summer School
Paris, June 14th 2018

Background

- *environmental sciences, early 80's;*
- *complex numerical and analytical models, based on first principles, conservation laws, ... ;*
- *ill-defined parameters, competing model structures
(different constitutive equations, different types of process considered, spatial/temporal resolution, ...)*
- *need to establish magnitude and sources of prediction uncertainty;*
- *Monte Carlo simulation analyses.*

MC filtering and RSA

Two tasks for RSA:

qualitative definition of the system behaviour

- [a set of constraints: thresholds, ceilings, time bounds based on available information on the system];

binary classification of model outputs based on the specified behaviour definition.

- [qualifies a simulation as **behaviour** (B) if the model output lies within constraints, **non-behaviour** (\bar{B}) otherwise]

MC filtering and RSA

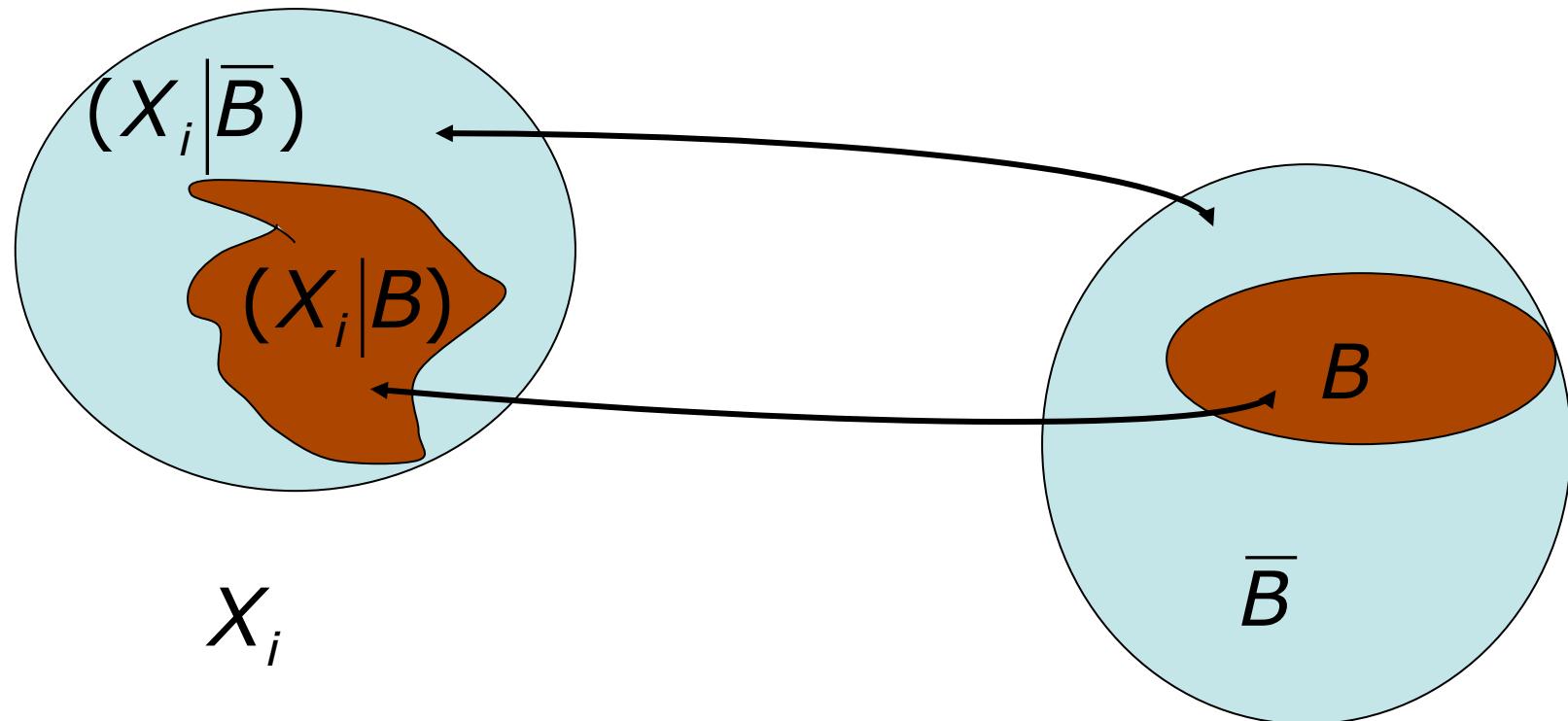
*Take a MC sample of size N of the **input factors** x_i ($i=1, \dots, k$), from ranges or prior distributions reflecting uncertainty in parameters and model constituent hypotheses.*

Each Monte Carlo simulation is associated to a vector of values of the input factors.

Classifying simulations as either B or \bar{B} , a set of binary elements is defined allowing to distinguish two sub-samples for each x_i ($n+m=N$):

size $m(x_i | B)$ size $n(x_i | \bar{B})$

MC filtering and RSA



MC filtering and RSA

The Kolmogorov-Smirnov (nonparametric) two-sample test (two-sided version) is performed for each factor *independently*

$$H_0 : f_m(x_i | B) = f_n(x_i | \bar{B})$$

$$H_1 : f_m(x_i | B) \neq f_n(x_i | \bar{B})$$

Under the null the two samples come from the same distribution.

Test statistic: $d_{m,n}(x_i) = \sup_y \|F_m(x_i | B) - F_n(x_i | \bar{B})\|$

F are marginal cumulative probability functions,
 f are probability density functions

MC filtering and RSA

At what significance level α does the computed value of $d_{m,n}$ determine the rejection of H_0 ?

P is the probability to reject H_0 when it is true (aka P-value i.e. to recognise a factor as important when it is not).

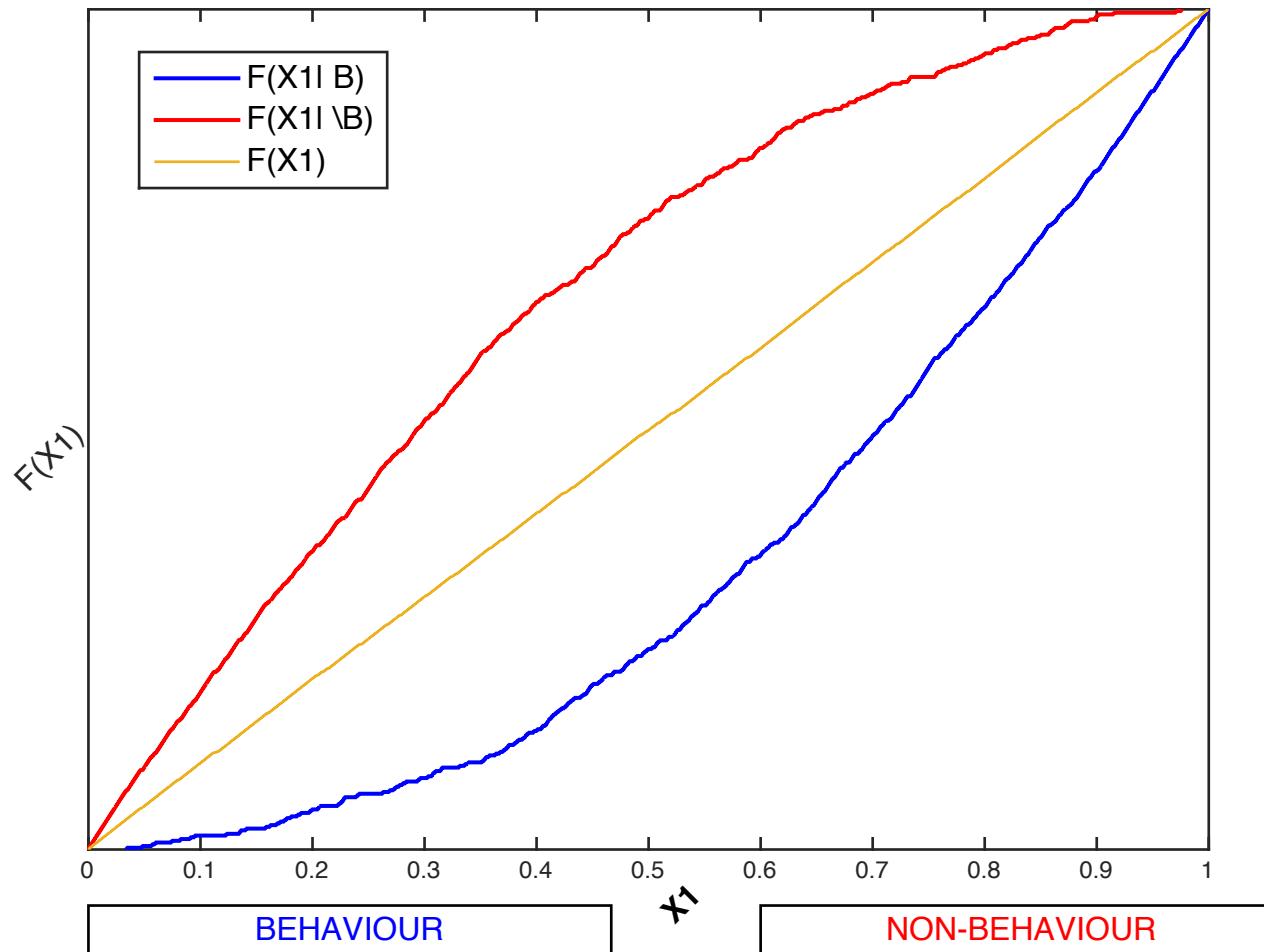
The importance of each parameter is inversely related to this probability.

MC filtering and RSA

Derive the critical level D_α (Smirnov statistics) above which the computed value of $d_{m,n}$ determines the rejection of H_0 (the smaller α , the higher D_α).

That is, if $d_{m,n} > D_\alpha$, then H_0 is rejected at significance level α (or $P\text{-value} < \alpha$).

MC filtering and RSA



Mapping unique solution: example (1)

Phillips curve:

GDP y , inflation π , output gap c , p potential, a_p
productivity shock, a_c demand shock, a_i cost push
shock

$$y_t = p_t + c_t$$

$$p_t = p_{t-1} + a_{p,t}$$

$$c_t = 2A \cos(2\pi / \tau) c_{t-1} - A^2 c_{t-2} + a_{c,t}$$

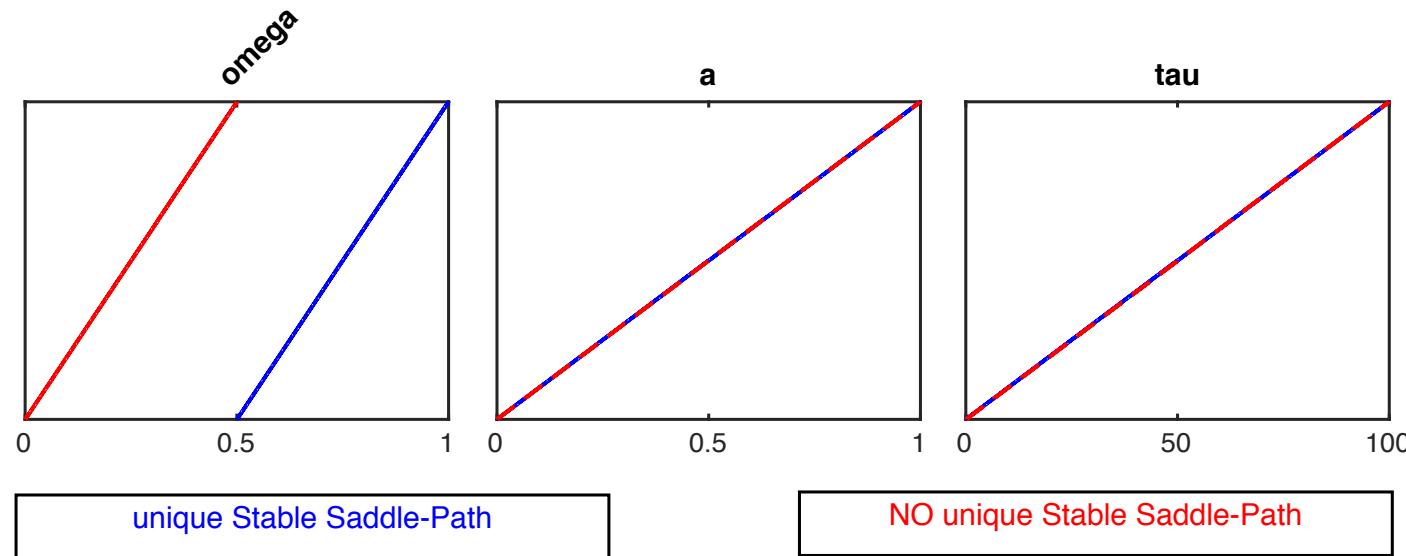
$$\pi_t = \omega \cdot \pi_{t-1} + (1 - \omega) E_t \pi_{t+1} + t_y c_t + a_{i,t}$$

Mapping unique solution: example (1)

Unique stable solution is obtained if:

$$\omega > 0.5$$

Mapping unique solution: example (1)



Mapping unique solution: example (2)

Phillips curve (extended version):

GDP y , inflation π , output gap c , p potential, a_p , productivity shock, a_c demand shock, a_i cost push shock

$$y_t = p_t + c_t$$

$$p_t = p_{t-1} + a_{p,t}$$

$$c_t = 2A \cos(2\pi / \tau) c_{t-1} - A^2 c_{t-2} + a_{c,t}$$

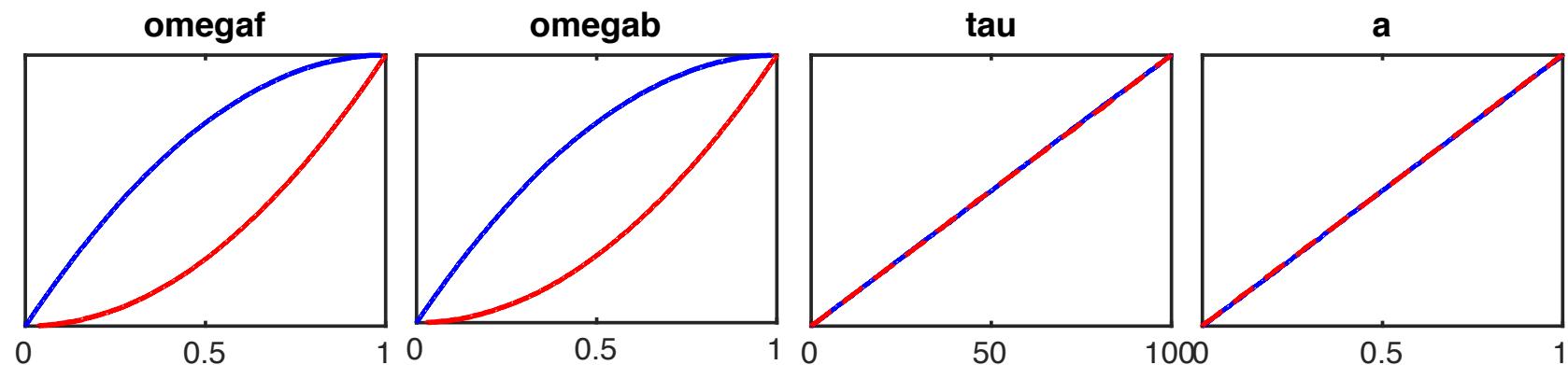
$$\pi_t = \omega_b \cdot \pi_{t-1} + \omega_f E_t \pi_{t+1} + t_y c_t + a_{i,t}$$

Mapping unique solution: example (2)

Unique stable solution is obtained if:

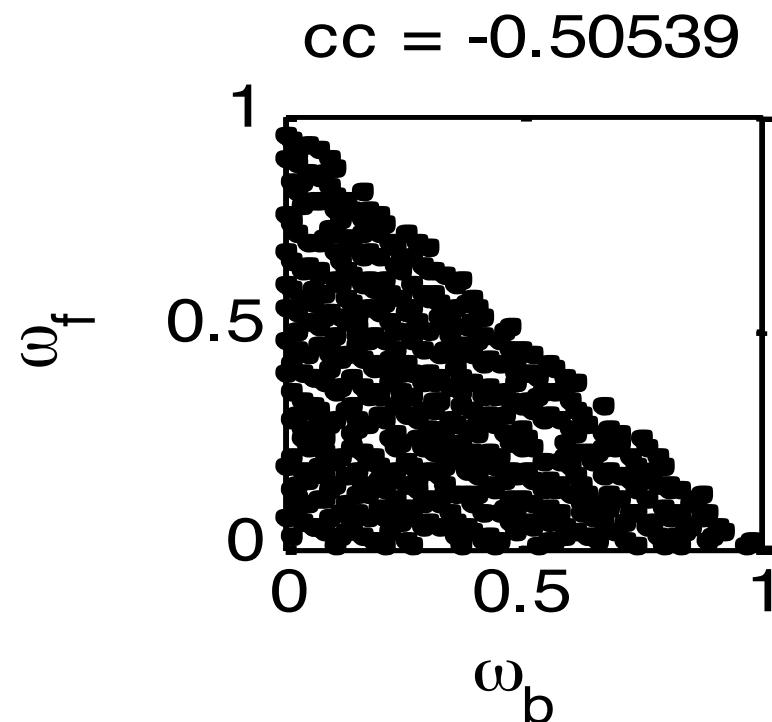
$$\omega_b + \omega_f < 1$$

Mapping unique solution: example (2)

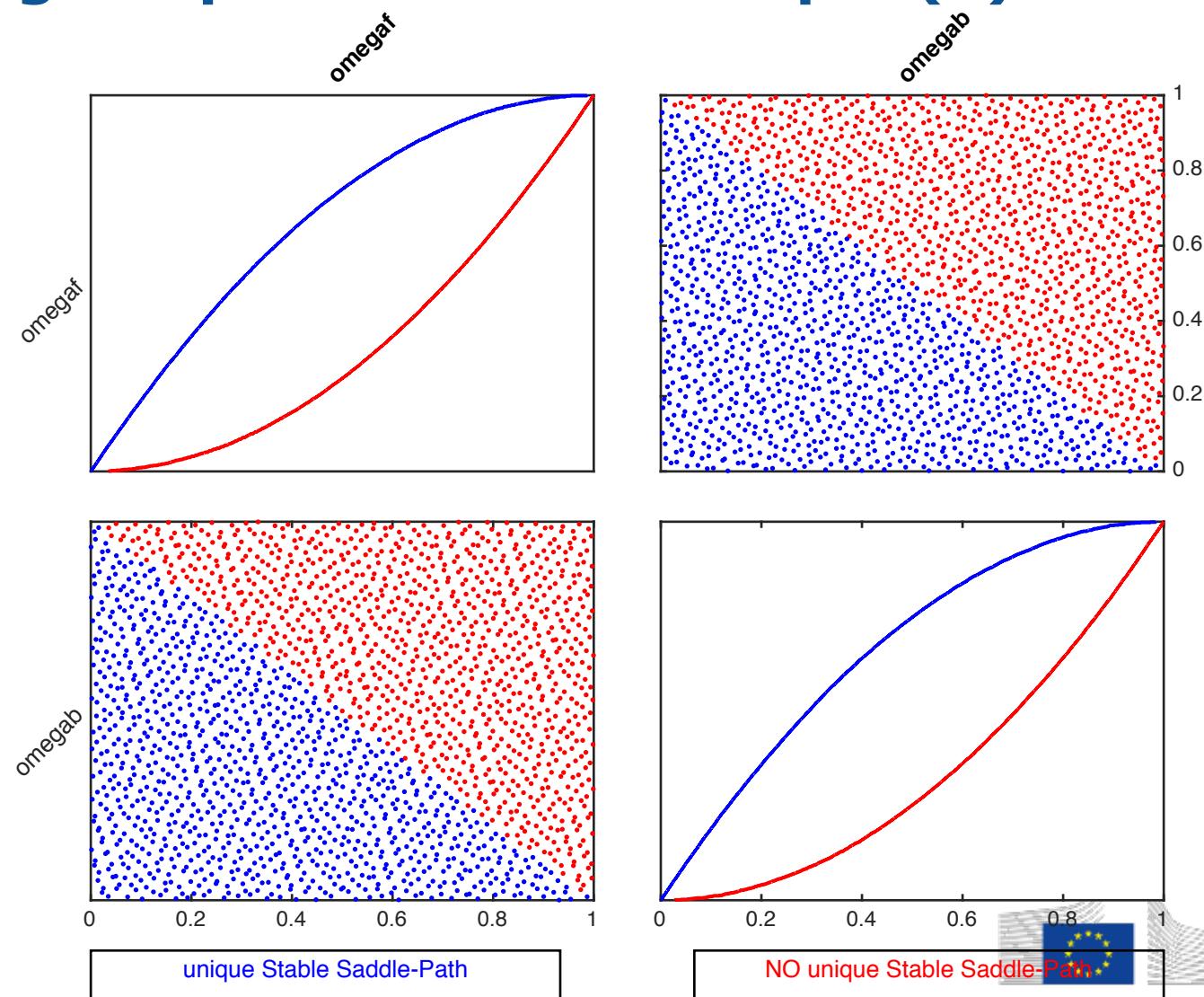


Mapping unique solution: example (2)

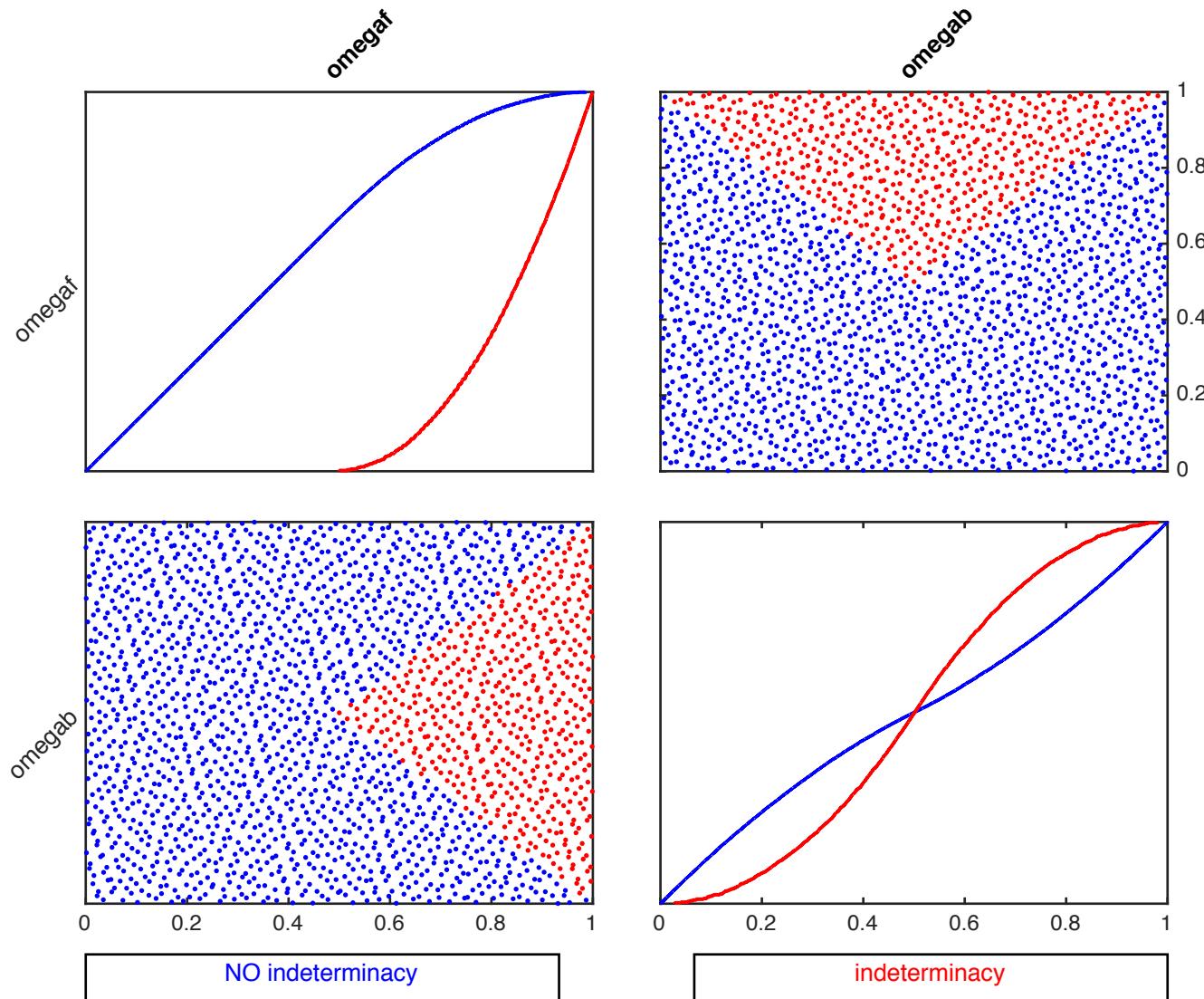
Bivariate analysis



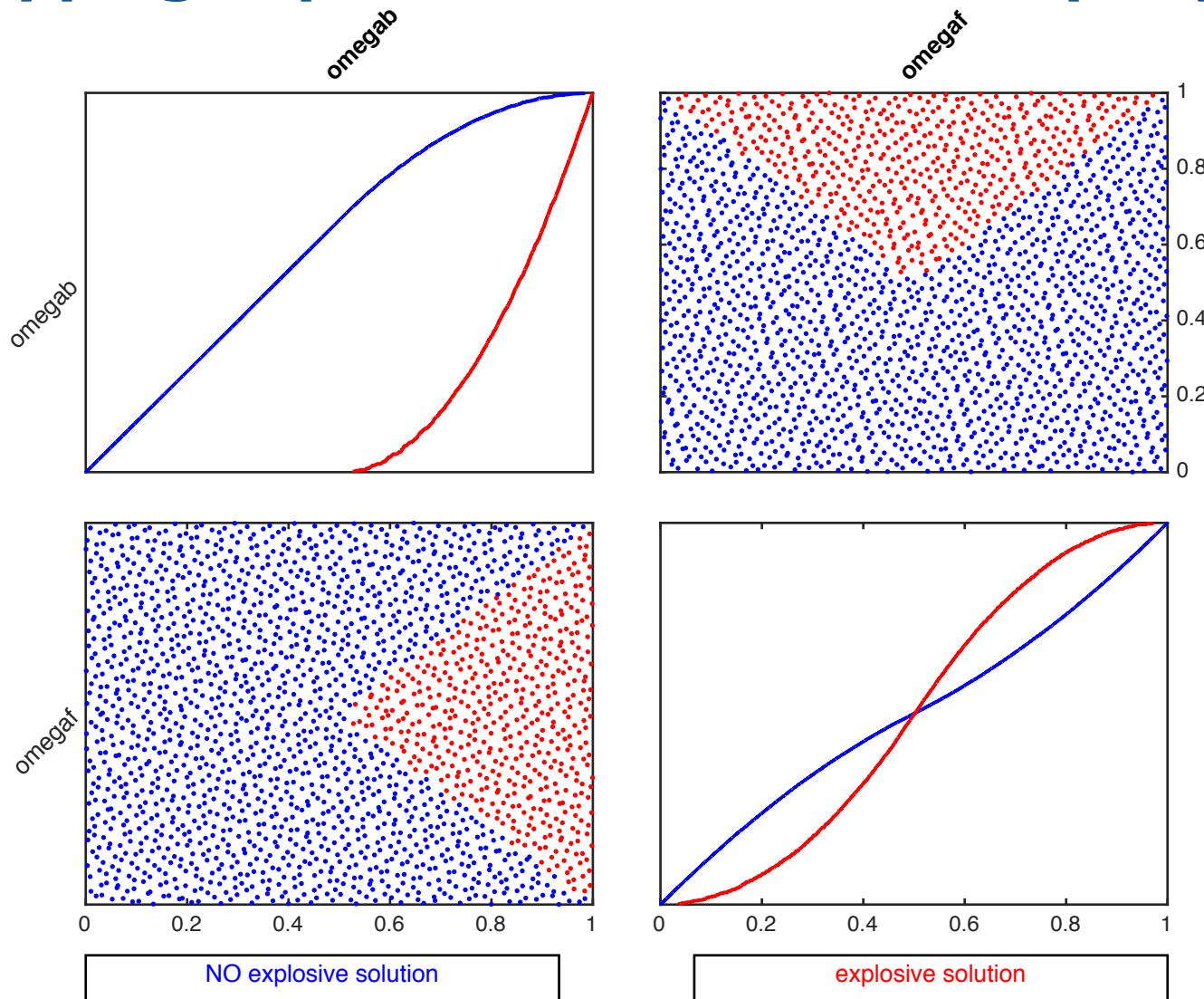
Mapping unique solution: example (2)



Mapping indeterminacy: example (2)



Mapping explosive solutions: example (2)



Mapping unique solution: example (3)

- *As the complexity of the model structure and its parameterization increases, it is not trivial to know a priori the set of model coefficients assuring stability;*

Mapping unique solution: example (3)

- *The Lubik Schorfheide (2005) model*
- *12 parameters;*

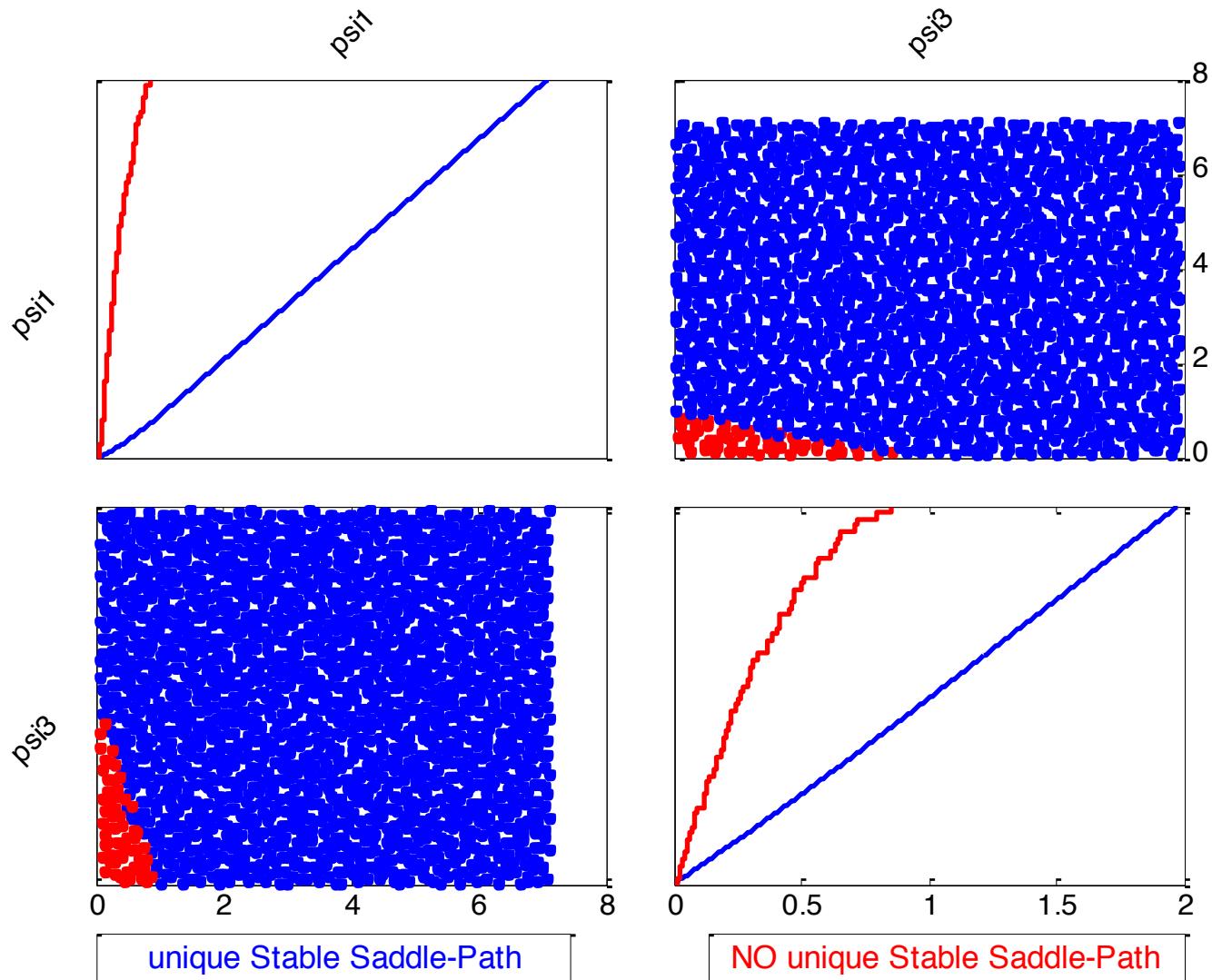
Lubik Schorfheide model (2005)

$$\begin{aligned}y_t &= E_t y_{t+1} - [\tau + \alpha(2 - \alpha)(1 - \tau)](R_t - E_t \pi_{t+1}) \\&\quad - \alpha[\tau + \alpha(2 - \alpha)(1 - \tau)]E_t \Delta q_{t+1} - \alpha(2 - \alpha)\frac{1 - \tau}{\tau}\Delta y_{t+1}^* - E_t z_{t+1} \\\\pi_t &= \beta E_t \pi_{t+1} + \alpha \beta E_t \Delta q_{t+1} - \alpha \Delta q_t + \frac{k}{\tau + \alpha(2 - \alpha)(1 - \tau)}(y_t - \bar{y}_t) \\\\pi_t &= \Delta e_t + (1 - \alpha)\Delta q_t + \pi_t^* \\\\mathcal{R}_t &= \rho_R R_{t-1} + (1 - \rho_R)(\psi_1 \pi_t + \psi_2(y_t - \bar{y}_t) + \psi_3 \Delta e_t) + e_{R,t} \\\\\Delta q_t &= \rho_q \Delta q_{t-1} + e_{q,t} \\\\mathcal{Y}_s &= \rho_{y^*} y_{t-1} + e_{y^*,t} \\\\pi_t^* &= \rho_{\pi^*} \pi_{t-1}^* + e_{\pi^*,t} \\\\mathcal{Z}_t &= \rho_z z_{t-1} + e_{z,t}\end{aligned}$$

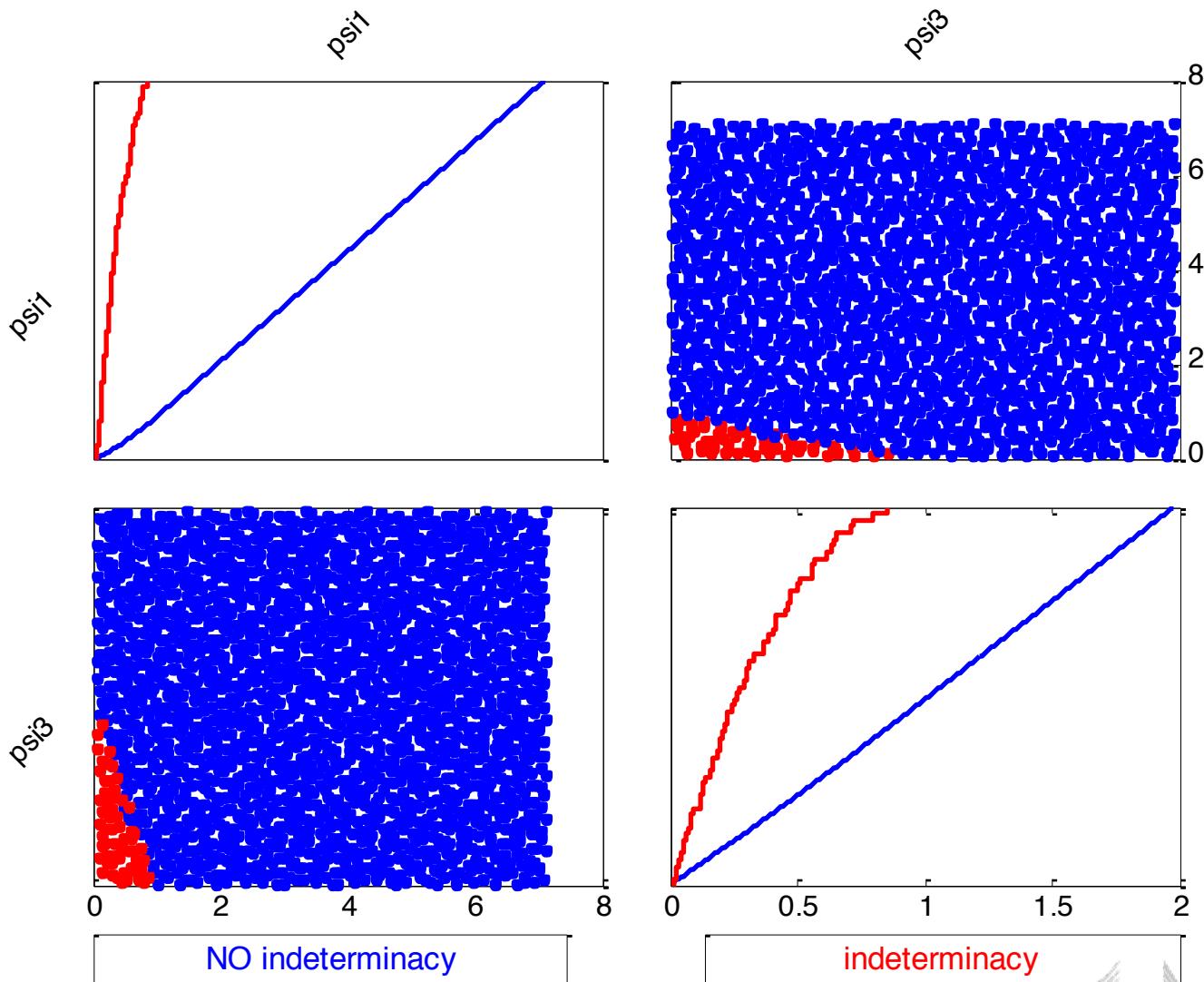
Lubik Schorfheide model (2005): priors

Name	Density	μ	σ
ψ_1	Gamma	1.5	0.5
ψ_2	Gamma	0.25	0.125
ψ_3	Gamma	0.25	0.125
ρ_R	Beta	0.5	0.2
α	Beta	0.3	0.1
r	Gamma	2.5	1
κ	Gamma	0.5	0.25
τ	Gamma	0.5	0.2
ρ_q	Gamma	0.4	0.2
ρ_z	Gamma	0.5	0.2
ρ_{y^*}	Gamma	0.8	0.1
ρ_{π^*}	Gamma	0.7	0.15
σ_R	InvGamma	1.2533	0.6551
σ_q	InvGamma	2.5066	1.3103
σ_z	InvGamma	1.2533	0.6551
σ_{y^*}	InvGamma	1.2533	0.6551
σ_{π^*}	InvGamma	1.88	0.9827

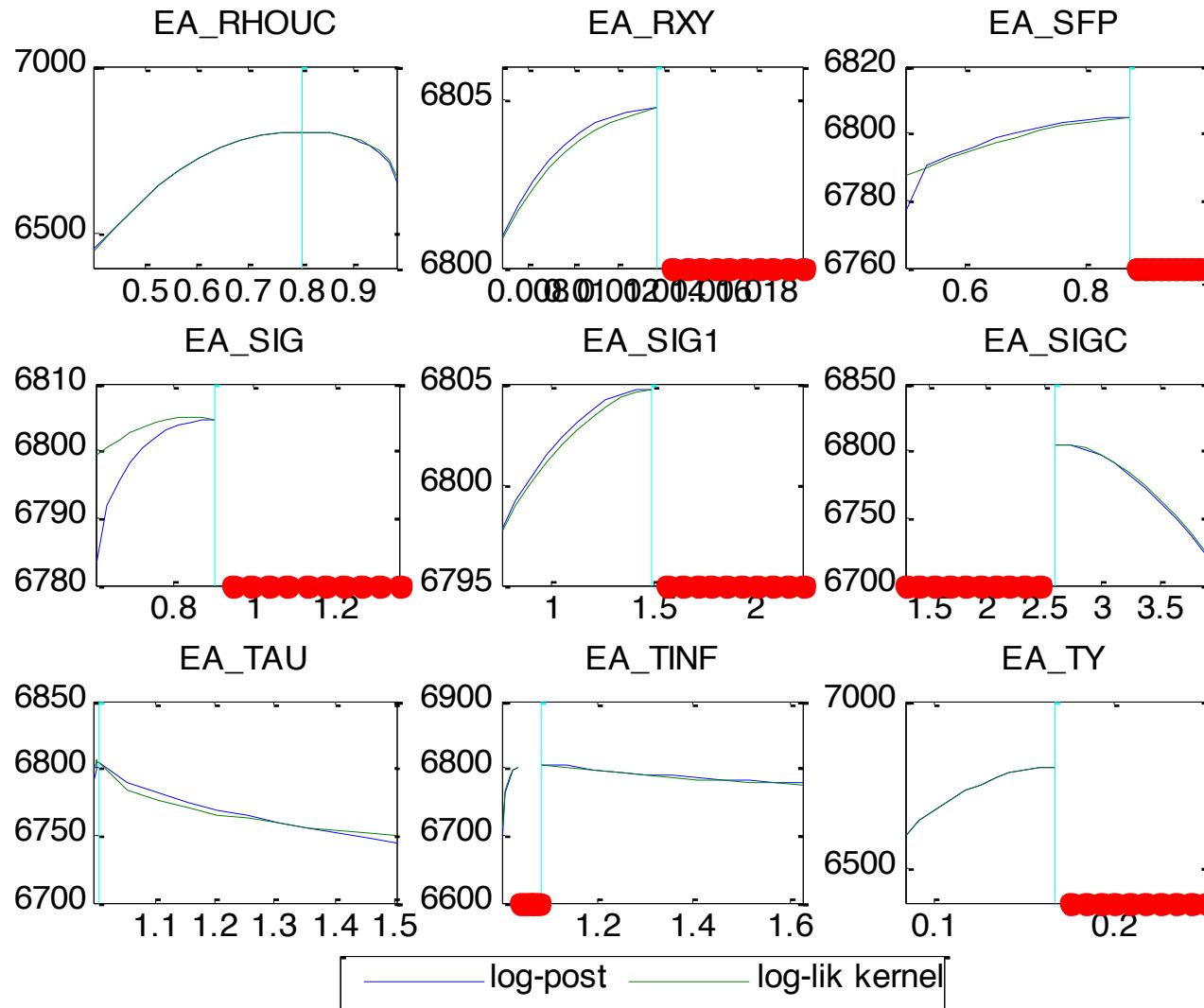
Lubik Schorfheide model (2005): unique solutions



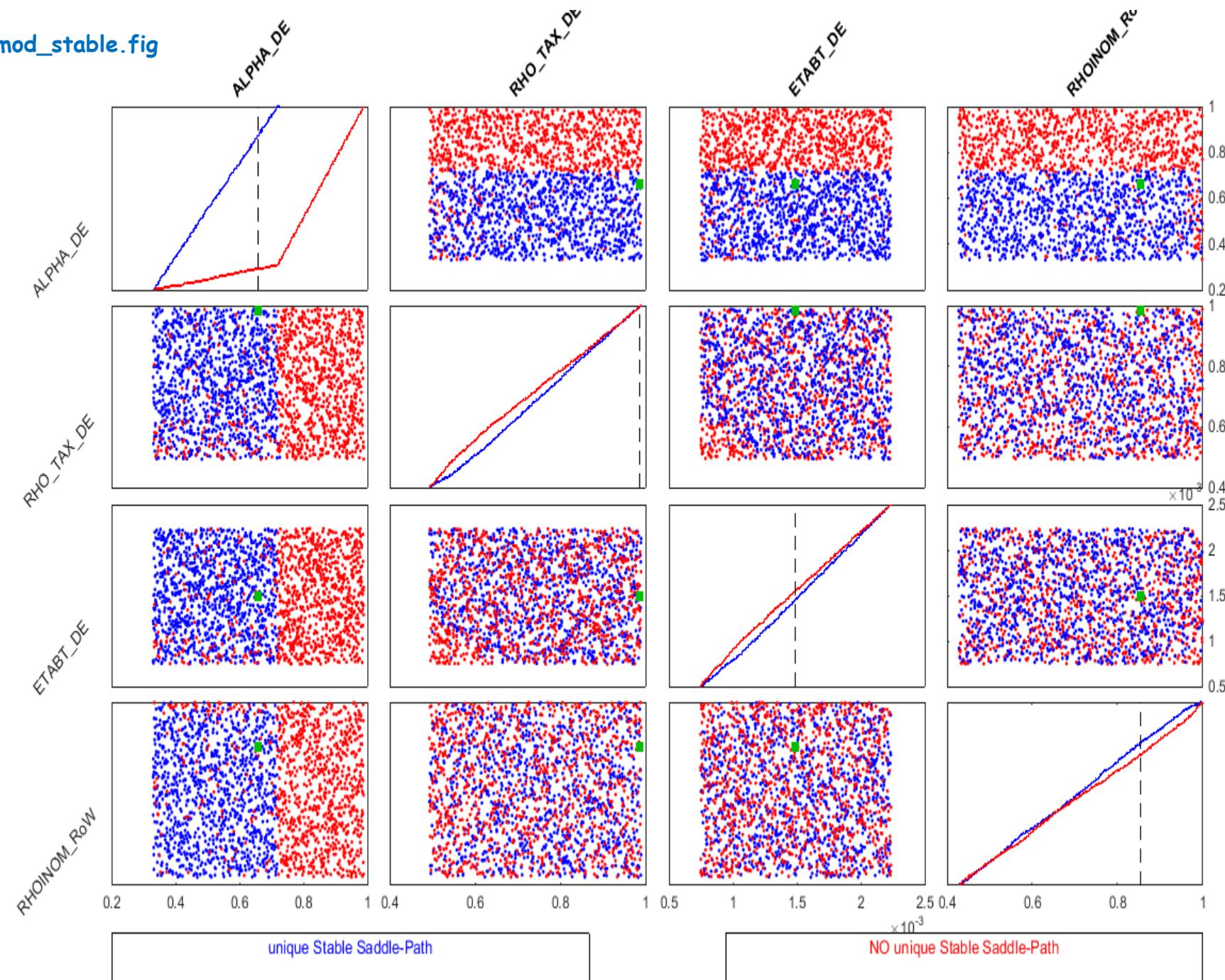
Lubik Schorfheide model (2005): indeterminacy



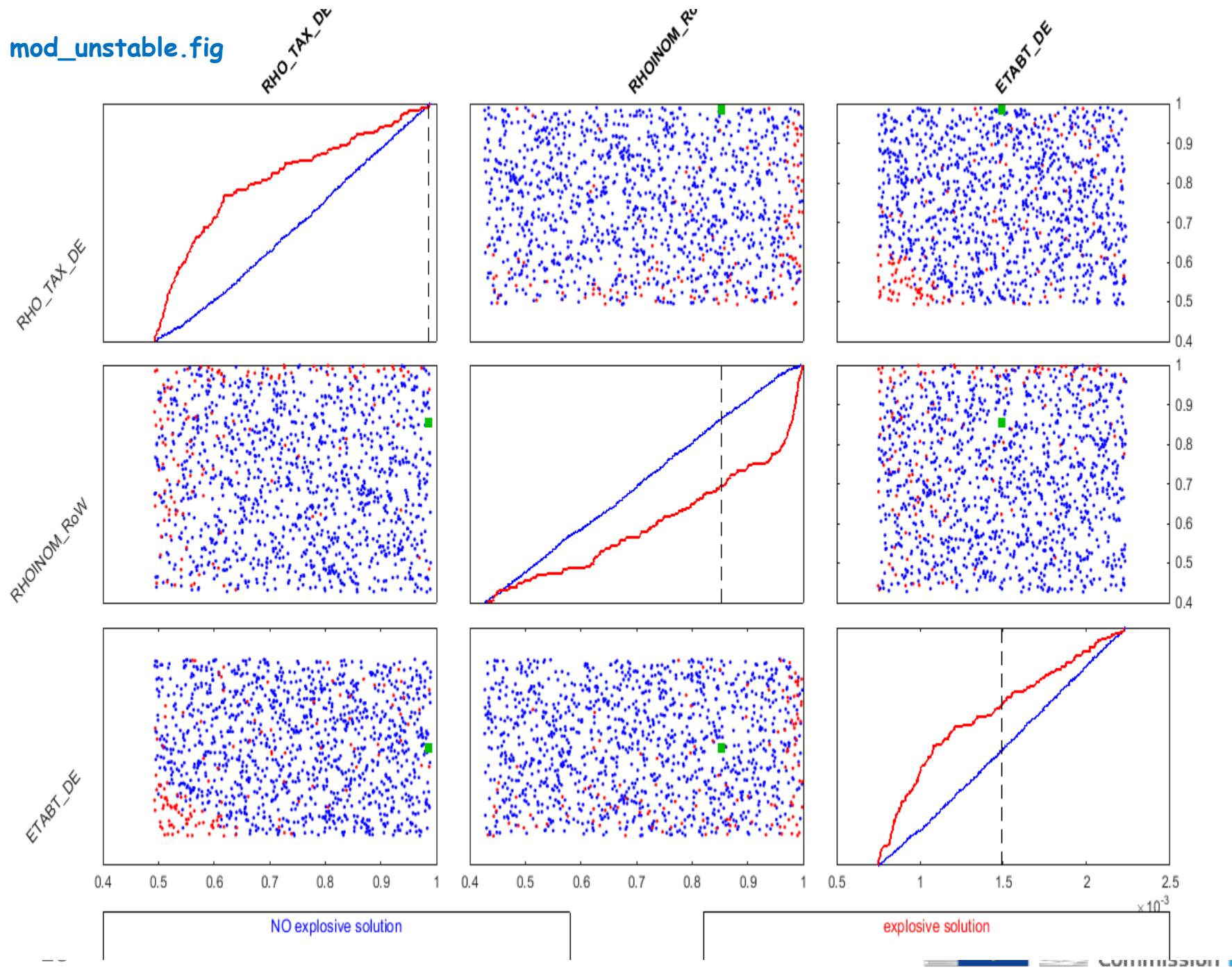
Diagnosing optimization problems: CLIFF



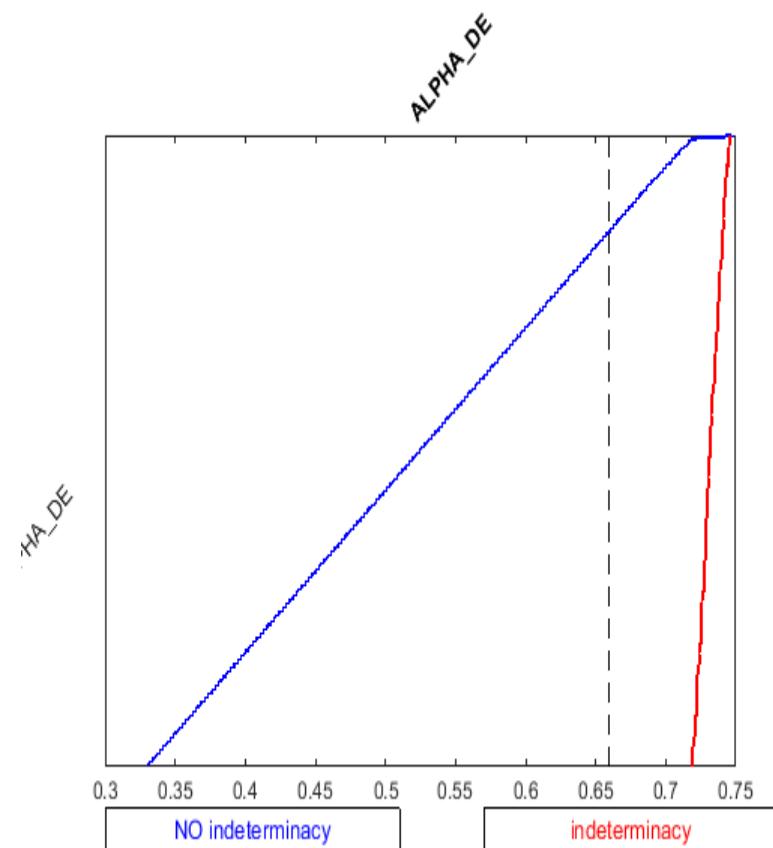
[mod_stable.fig](#)



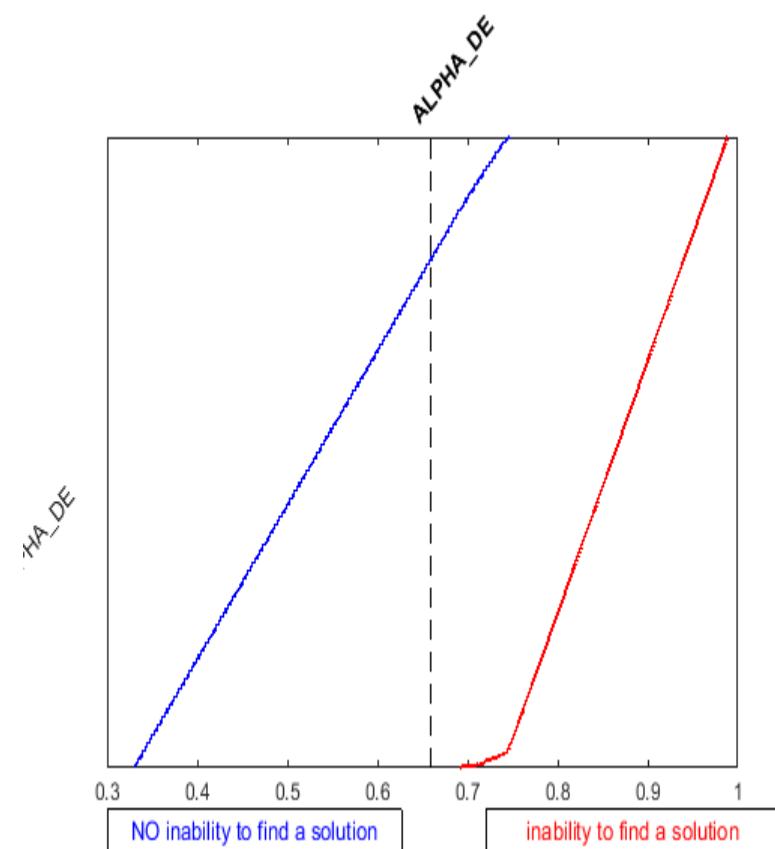
mod_unstable.fig



mod_indeterm.fig



mod_wrong.fig



Toolbox documentation

3.1 Sampling options

option name	default	description
Nsam	2048	Size of MC sample
ilptau	1	1 = use LP_τ quasi-Monte Carlo 0 = use LHS Monte Carlo
pprior	1	1 = sample from prior distributions 0 = sample from multivariate normal $N(\hat{\theta}, \Sigma)$, $\hat{\theta}$ is posterior mode $\Sigma = H^{-1}$, H is Hessian at the mode
prior_range	1	1 = sample <i>uniformly</i> from prior ranges 0 = sample from prior distributions:
morris	0	0 = no Morris sampling for screening 1 = Morris sampling for screening
morris_nliv	6	number of levels in Morris design
morris_ntra	20	number of trajectories in Morris design
ppost	0	0 = don't use Metropolis posterior sample 1 = use Metropolis posterior sample: this overrides any other sampling option!
neighborhood_width	[]	δ (real number > 0) uniform sample in the neighborhood of the posterior mode $\hat{\theta}$ interval width: $\hat{\theta}(1 \pm \delta)$

Also requires mode_file option

Toolbox documentation

3.2 Stability mapping

option name	default	description
stab	1	1 = perform stability mapping 0 = no stability mapping is performed
load_stab	0	0 = generate a new sample 1 = load a previously created sample
pvalue_corr	0.001	critical p-value for correlations ρ in filtered samples: plot couples of parameters with $p\text{-value} < \text{pvalue_corr}$
pvalue_ks	0.001	critical p-value for Smirnov statistics d : plot parameters with $p\text{-value} < \text{pvalue_ks}$
lik_init	1	1 = the model is stationary (unit roots are ‘explosive’) 3 = the model has unit roots (unit roots are ‘stable’)

Toolbox documentation

examples

```
dynare_sensitivity(prior_range=0, load_stab=0, nodisplay,  
graph_format=(eps,fig), lik_init=3, Nsam=512);
```

```
dynare_sensitivity(pprior=0, neighborhood_width=0.1,  
mode_file=quest3hlmr_mode, lik_init=3, nodisplay,  
graph_format=(eps,fig), Nsam=512);
```

IRF/Moment calibration

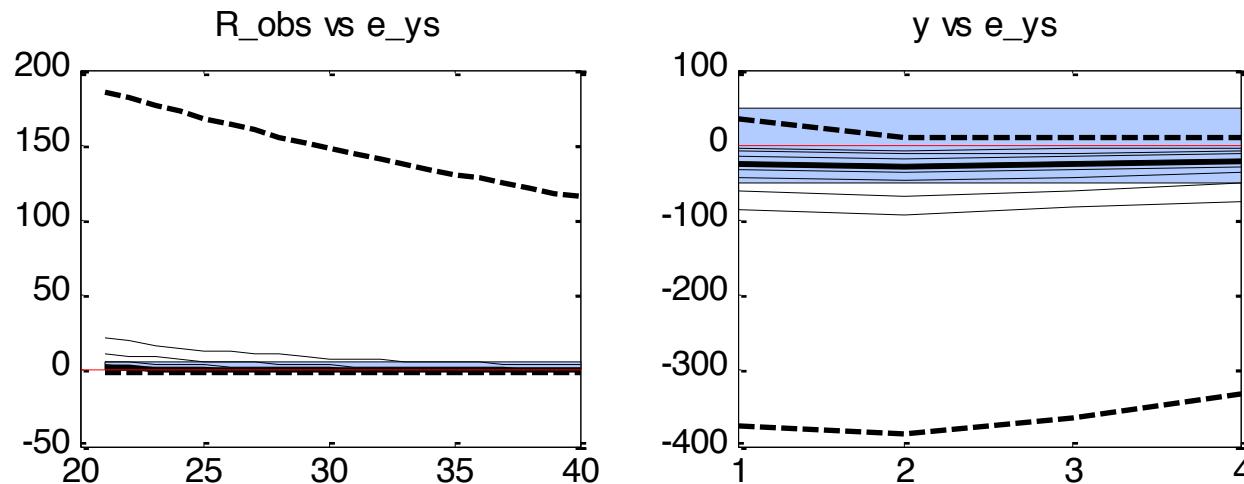
1) IRF calibration

- Restricting impulse responses (drivers)
- sign changes (oscillations)
- target benchmark model (model comparison)

2) Moment calibration (theoretical moments)

- Match key properties in the data (auto or cross correlation, sign restrictions)

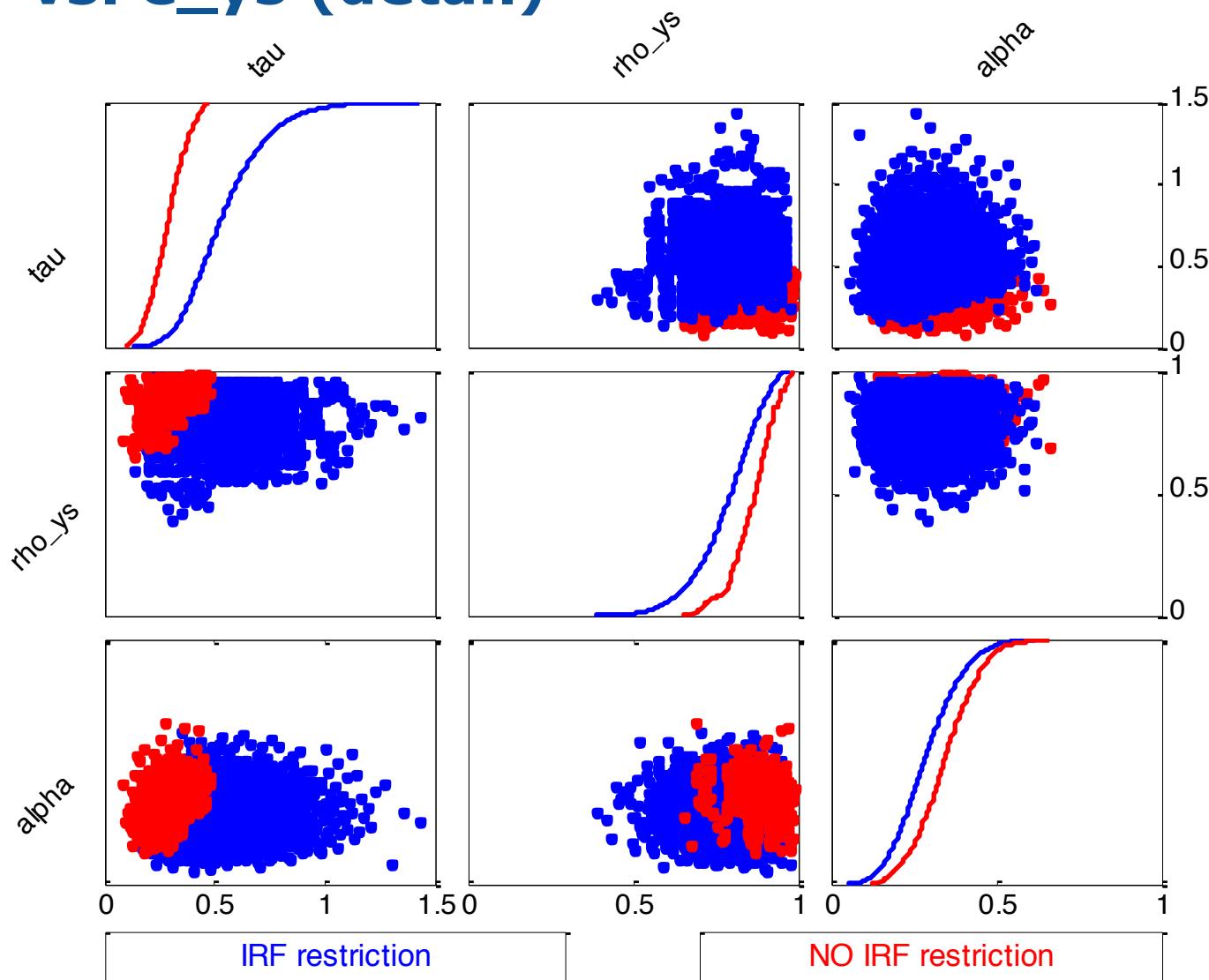
IRF calibration



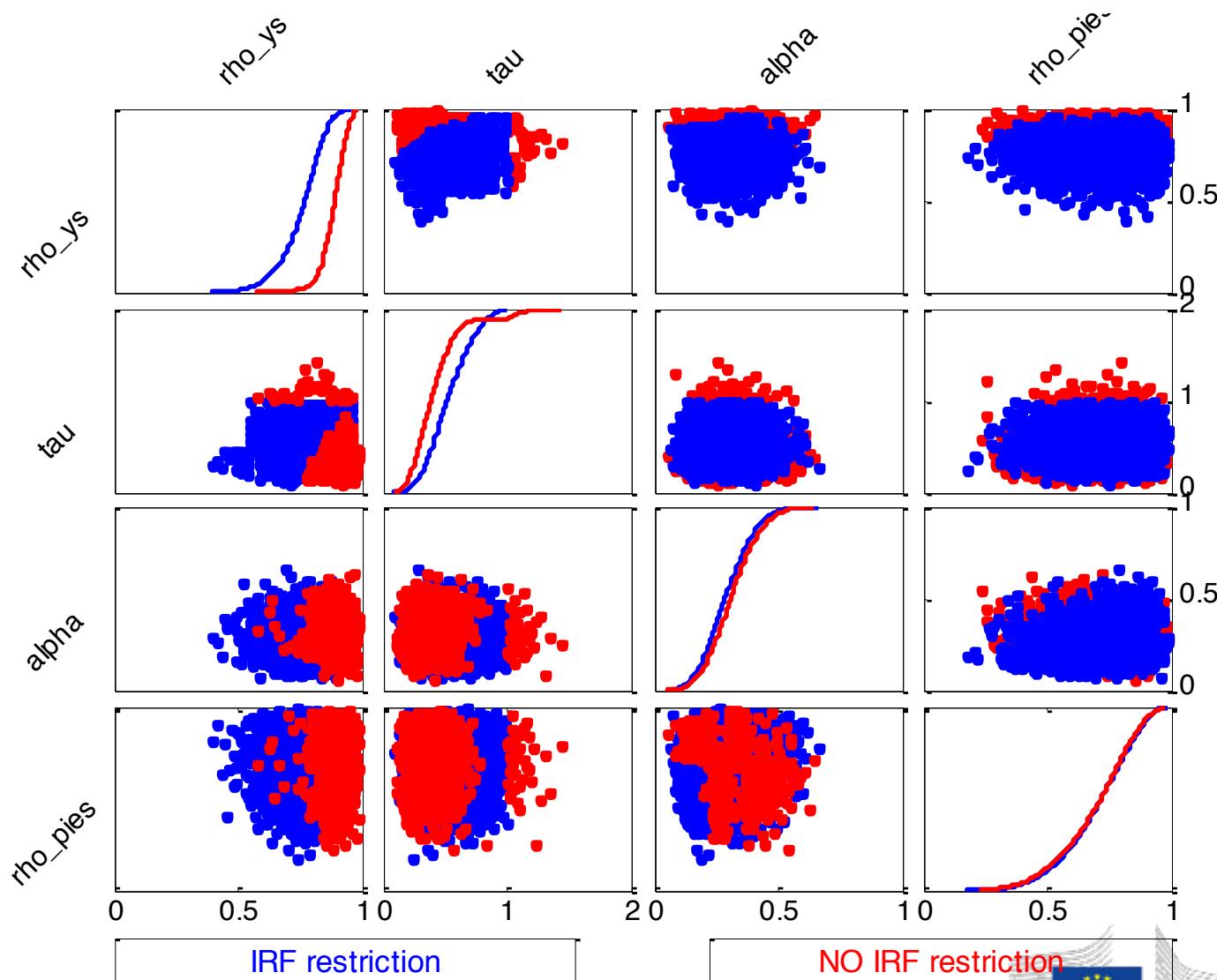
```
irf_calibration(relative_irf);
y(1:4), e_ys, [-50 50]; // [first year response, logical OR]
@#for ilag in 21:40
R_obs(@{ilag}), e_ys, [0 6]; // [response after 4th year to 10th year]
@#endfor
end;
```

logical AND

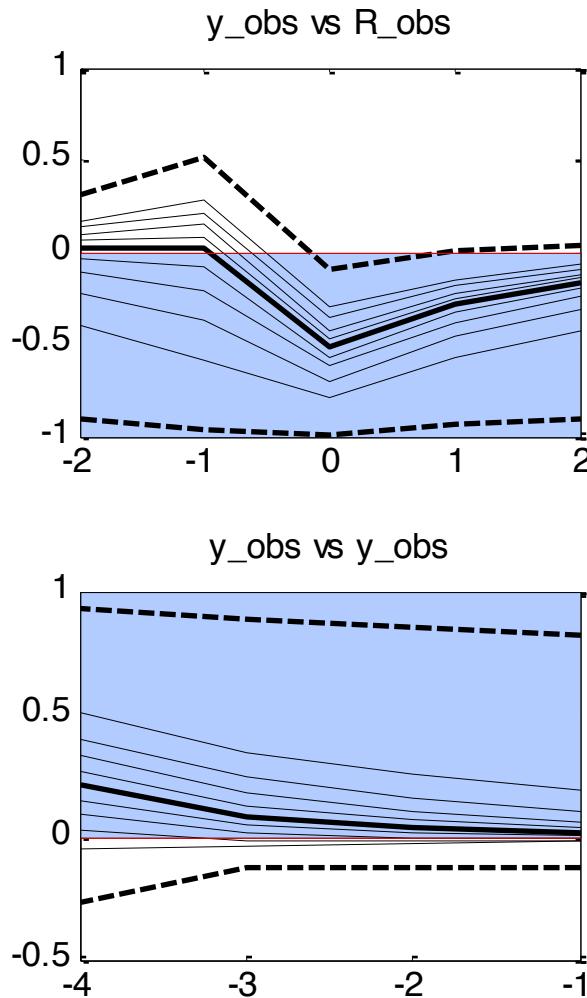
IRF y vs. e_ys (detail)



IRF R_obs vs. e_ys (detail)

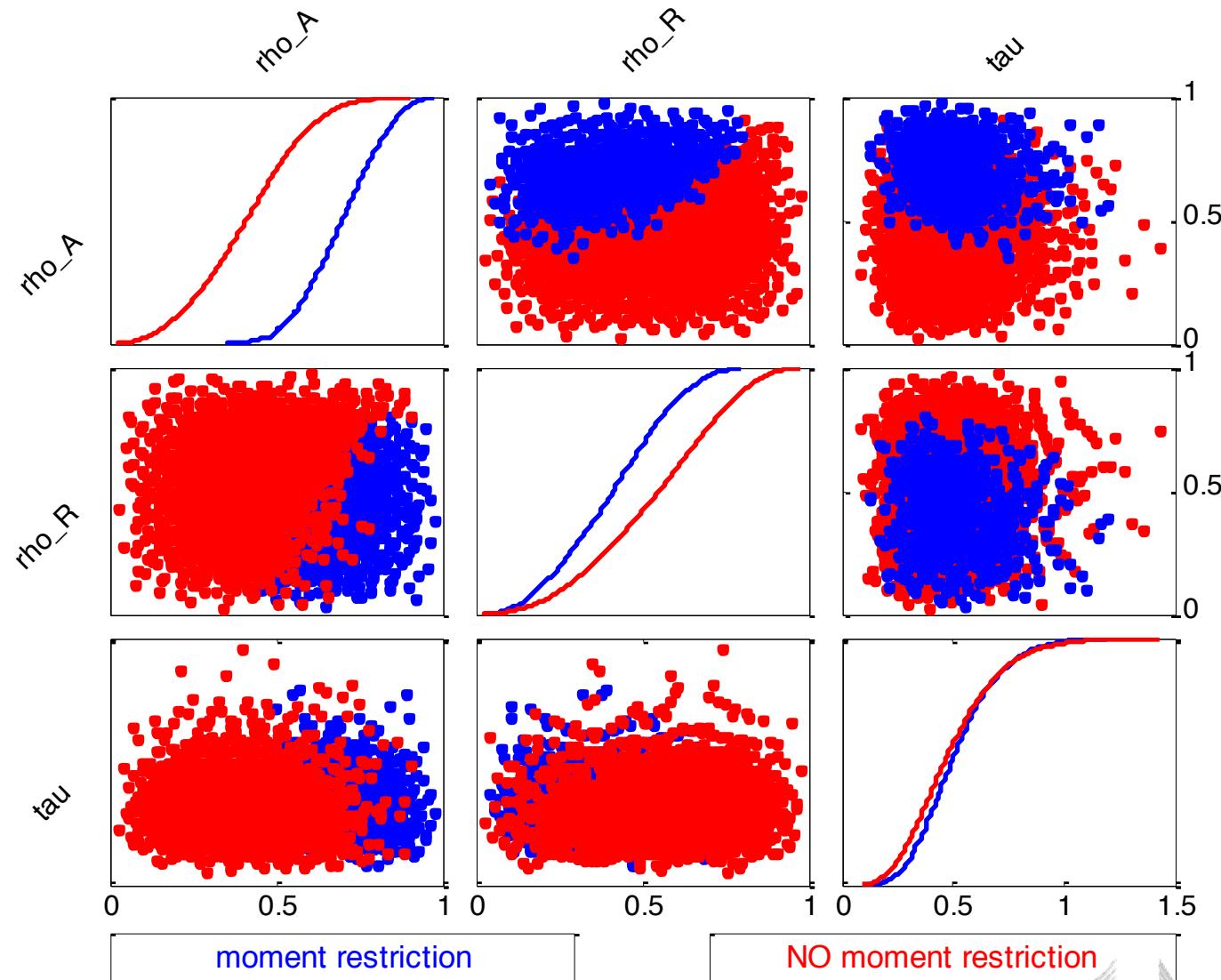


Moment calibration

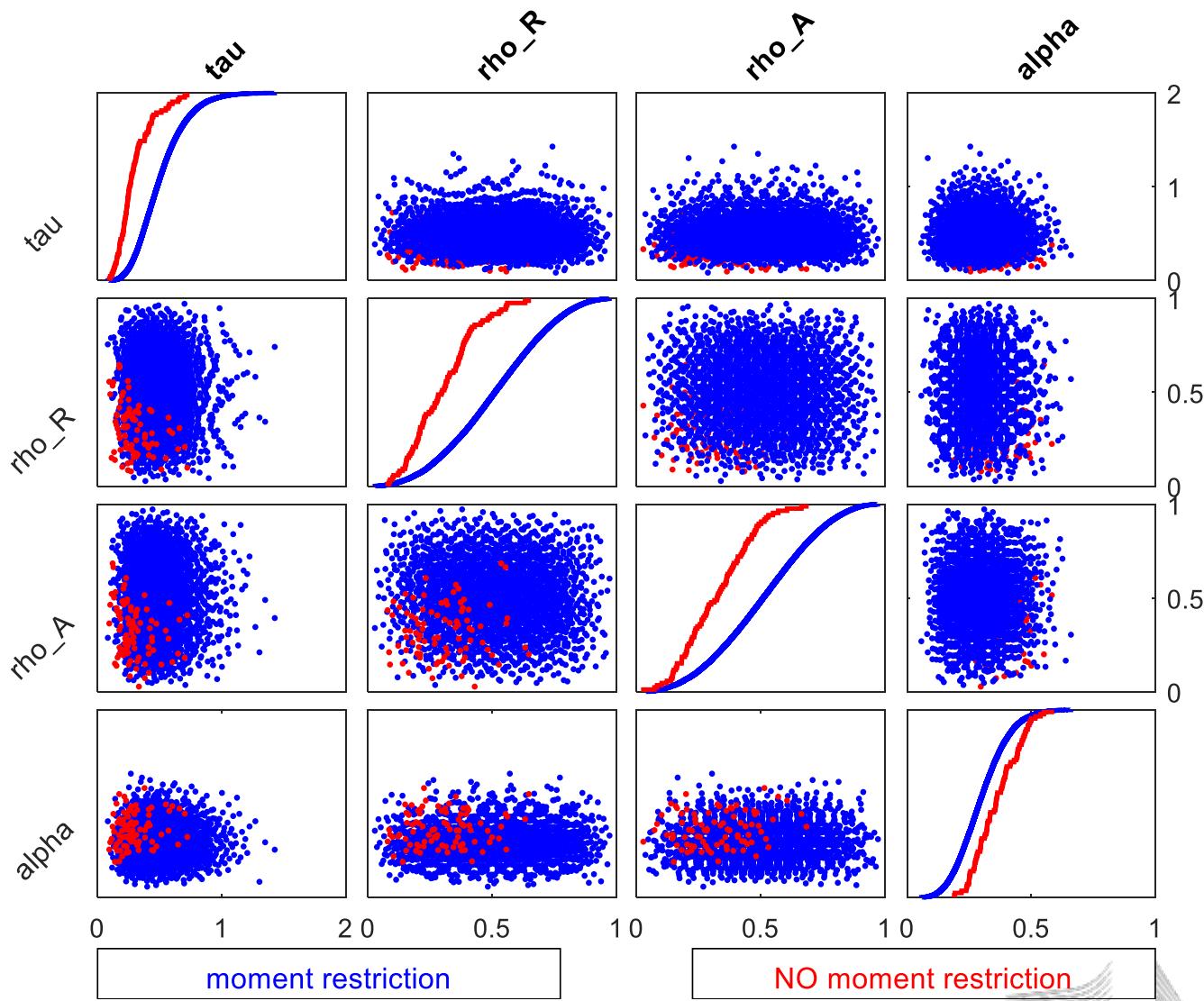


```
moment_calibration;  
y_obs,y_obs(-(1:4)), +; // [first year acf] OR  
@#for ilag in -2:2  
y_obs,R_obs(@{ilag}), -; // [ccf] AND  
@#endfor  
@#for ilag in -4:4  
y_obs,pie_obs(@{ilag}), -; // [ccf] AND  
@#endfor  
end;
```

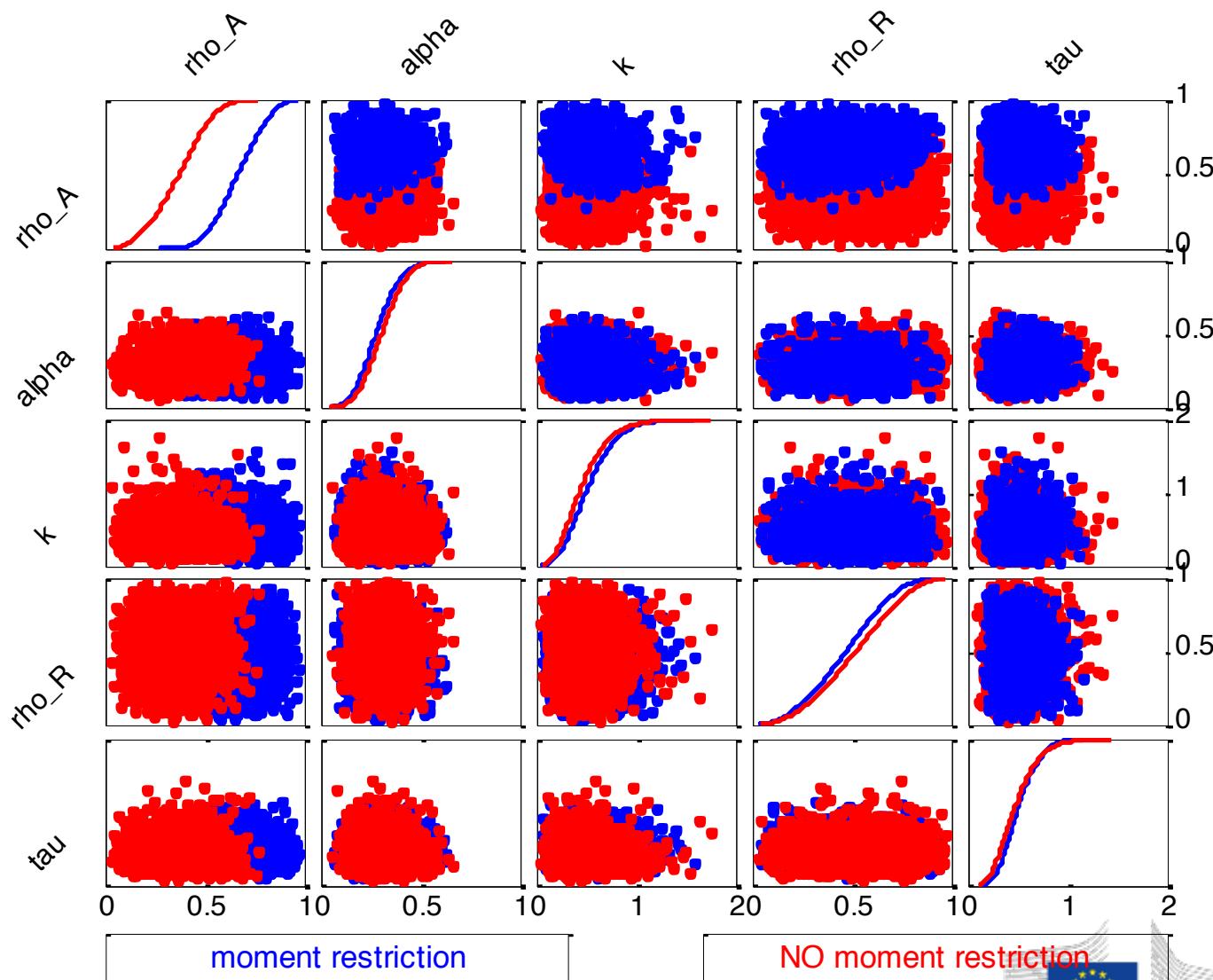
Moment y_{obs} vs. $\pi_{obs}(-4...4)$



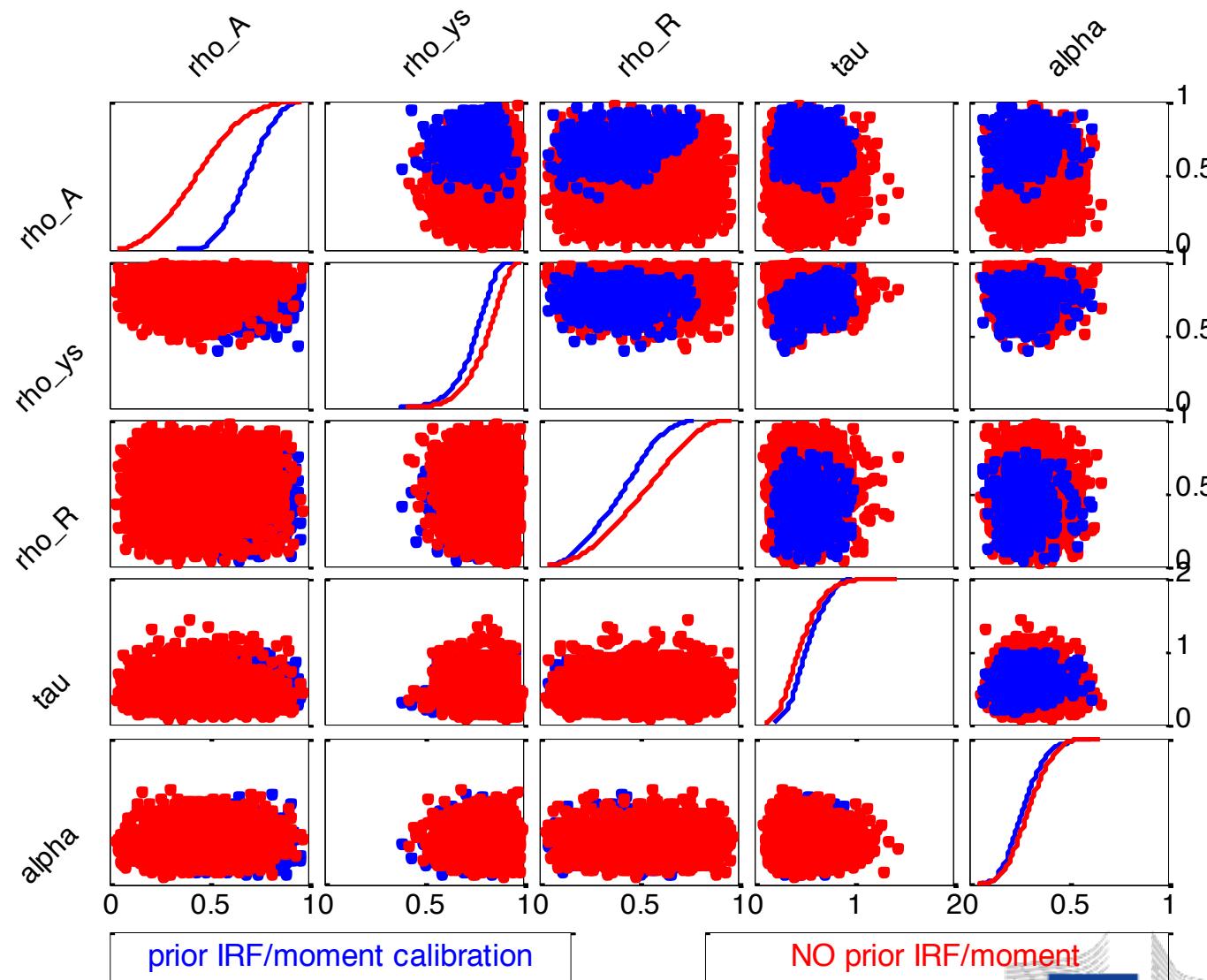
Moment y_{obs} vs. $\pi_{obs}(-1)$



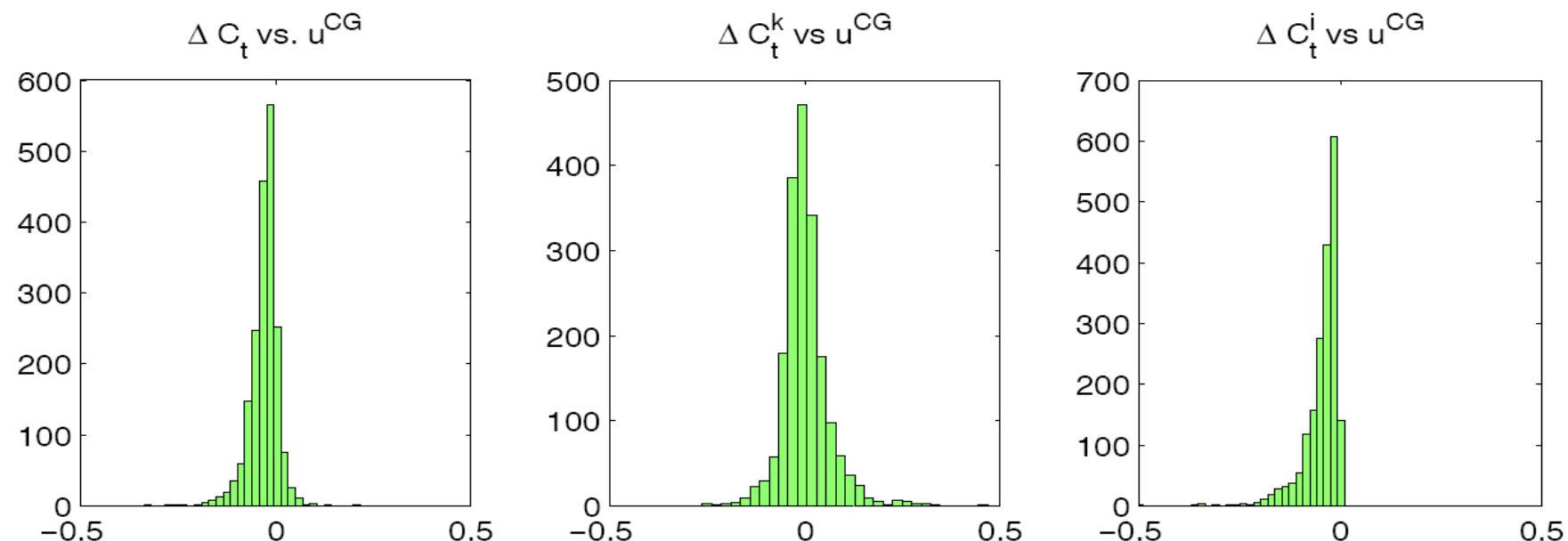
Moment y_{obs} vs. $R_{obs}(-2...2)$



IRF/Moment calibration

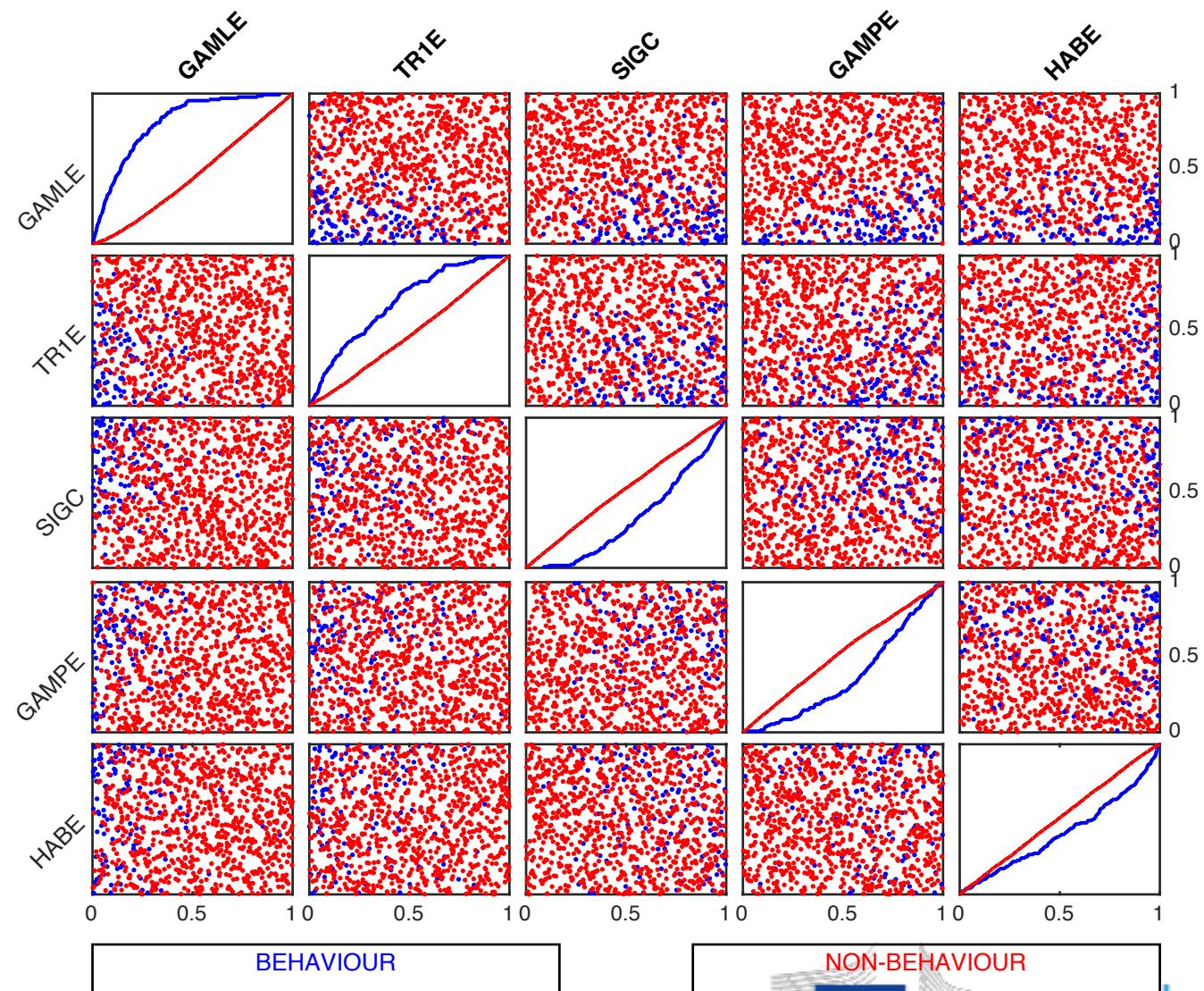


Ratto M, W. Roeger and J. in 't Veld (2009) QUEST III: consumption response vs. government spending shock

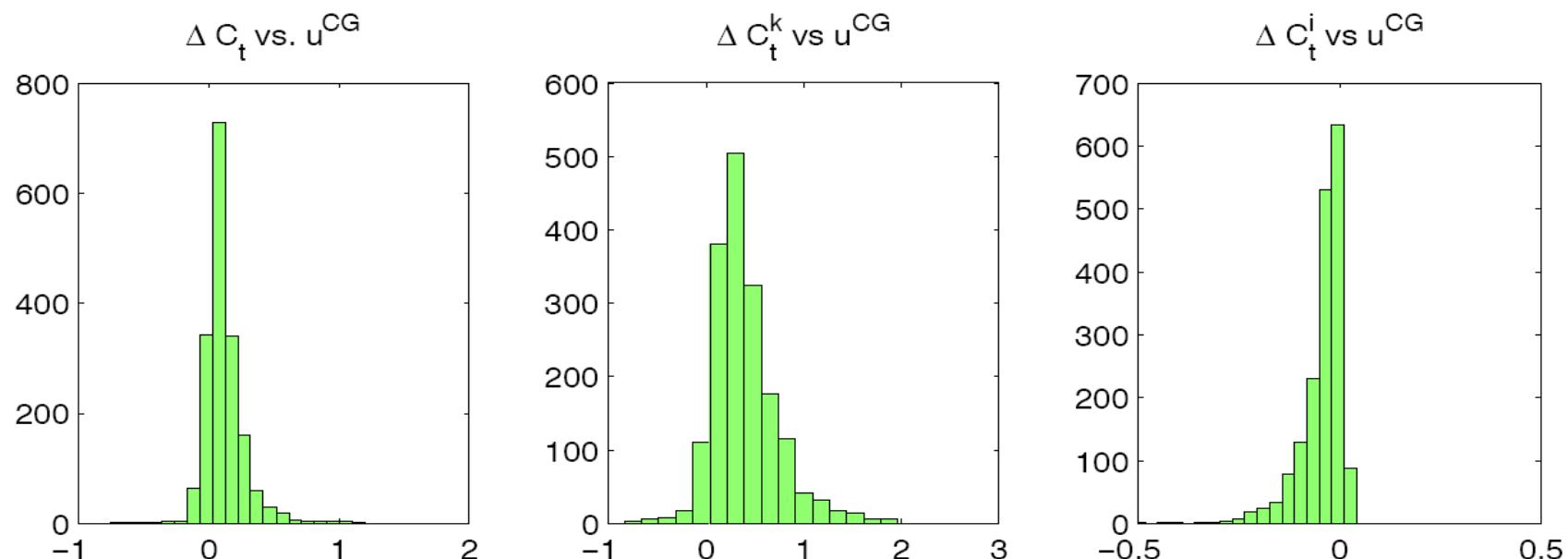


Prior distribution of the response of aggregate consumption of LC
and Ricardian HH versus government spending shock

Target
behavior:
crowding in of
C as in Gali et
al. 2007



Ratto M, W. Roeger and J. in 't Veld (2009) QUEST III: consumption response vs. government spending shock



Prior distribution of the response of aggregate consumption versus government spending eliminating labor demand adj. cost as in Gali *et al.* (2007)

Mapping the set of observables

How do parameters adjust to fit the multivariate set of observed covariates?

How do they change by adding/removing observables?

Are there trade-offs?

Mapping the set of observables

- *Sample structural coefficients from prior distributions (or ranges);*
OR
- *Use Metropolis posterior sample;*

Compute RMSE's of the 1-step ahead model predictions for each of the N observed series.

Mapping the set of observables

*N filtering rules:
defining as B the samples corresponding to the
best 10% RMSE's, \bar{B} otherwise.*

*For each parameter we get N ‘behavioural’ sets:
 $f_j(X_i|B); j = 1, \dots, N$*

Mapping the set of observables

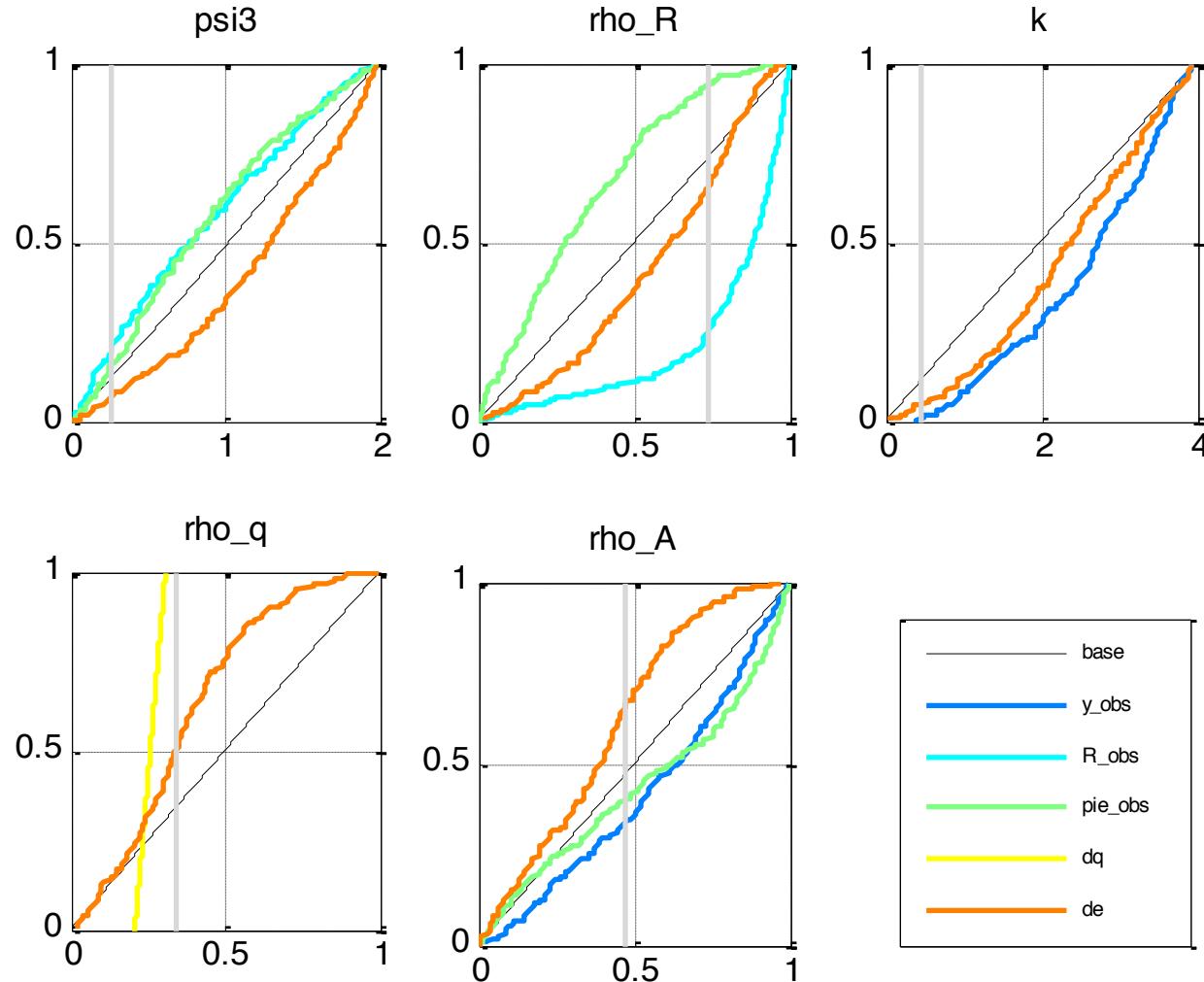
Trade off occurs when:

1) one structural parameter drives significantly the fit of more than one observed series;

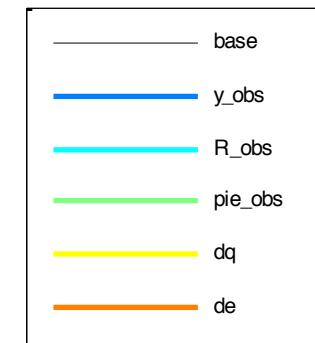
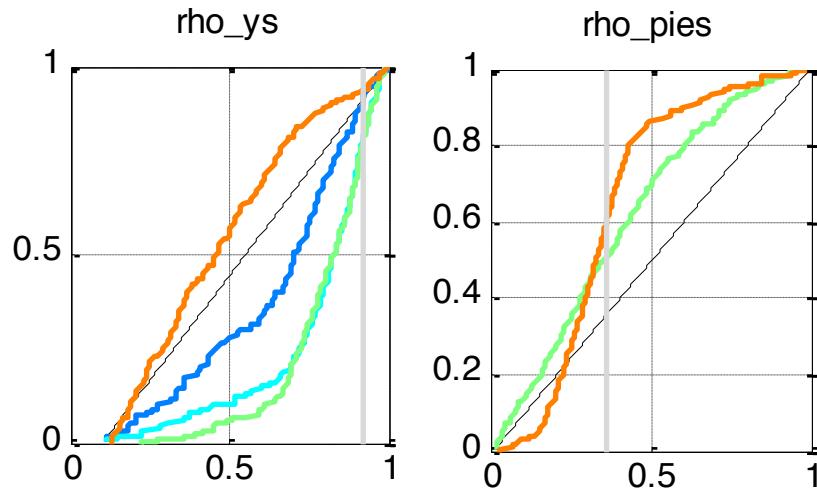
AND

2) $f_j(X_i|B) \neq f_k(X_i|B)$

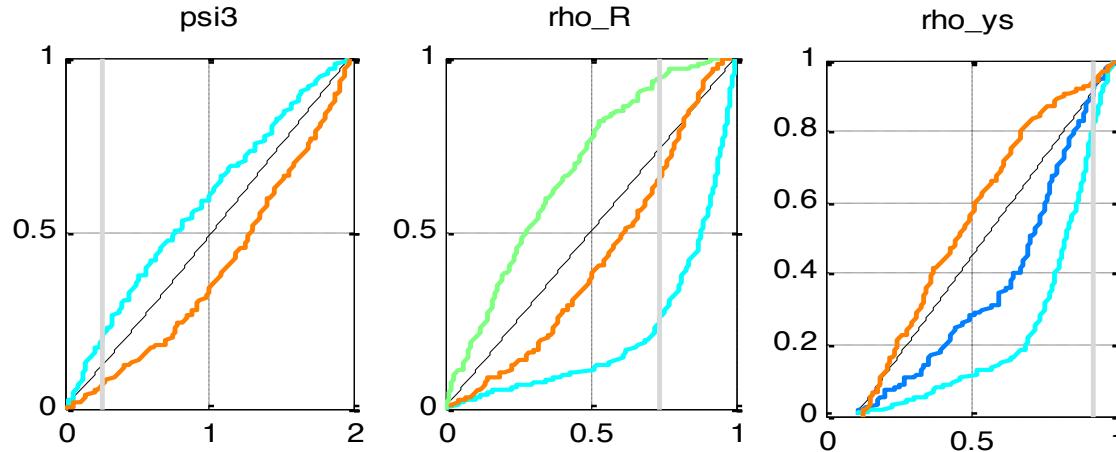
Mapping the fit: prior



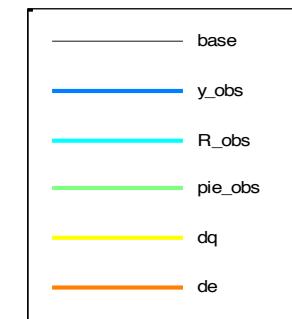
Mapping the fit: prior



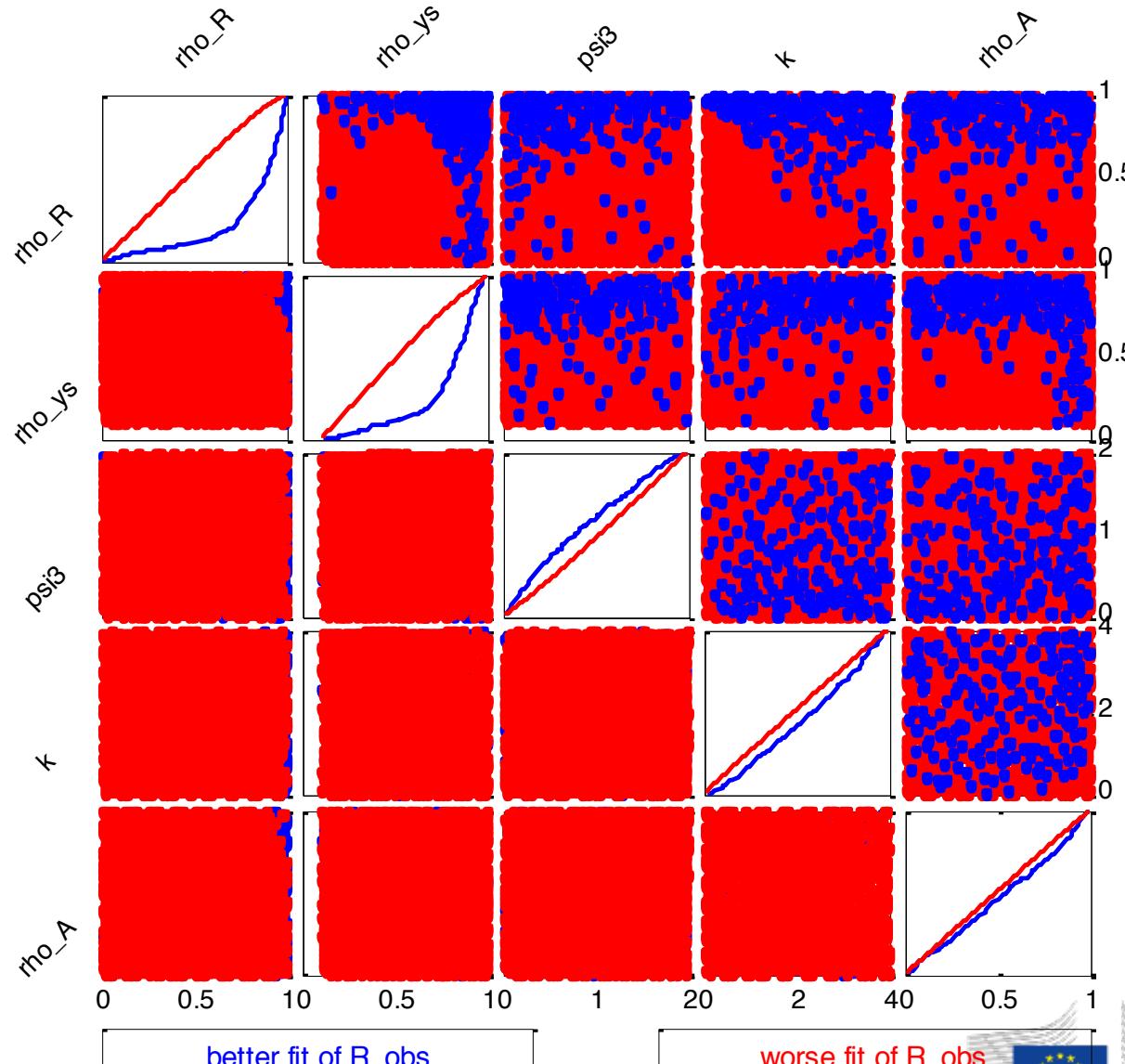
Mapping the fit by observable: prior



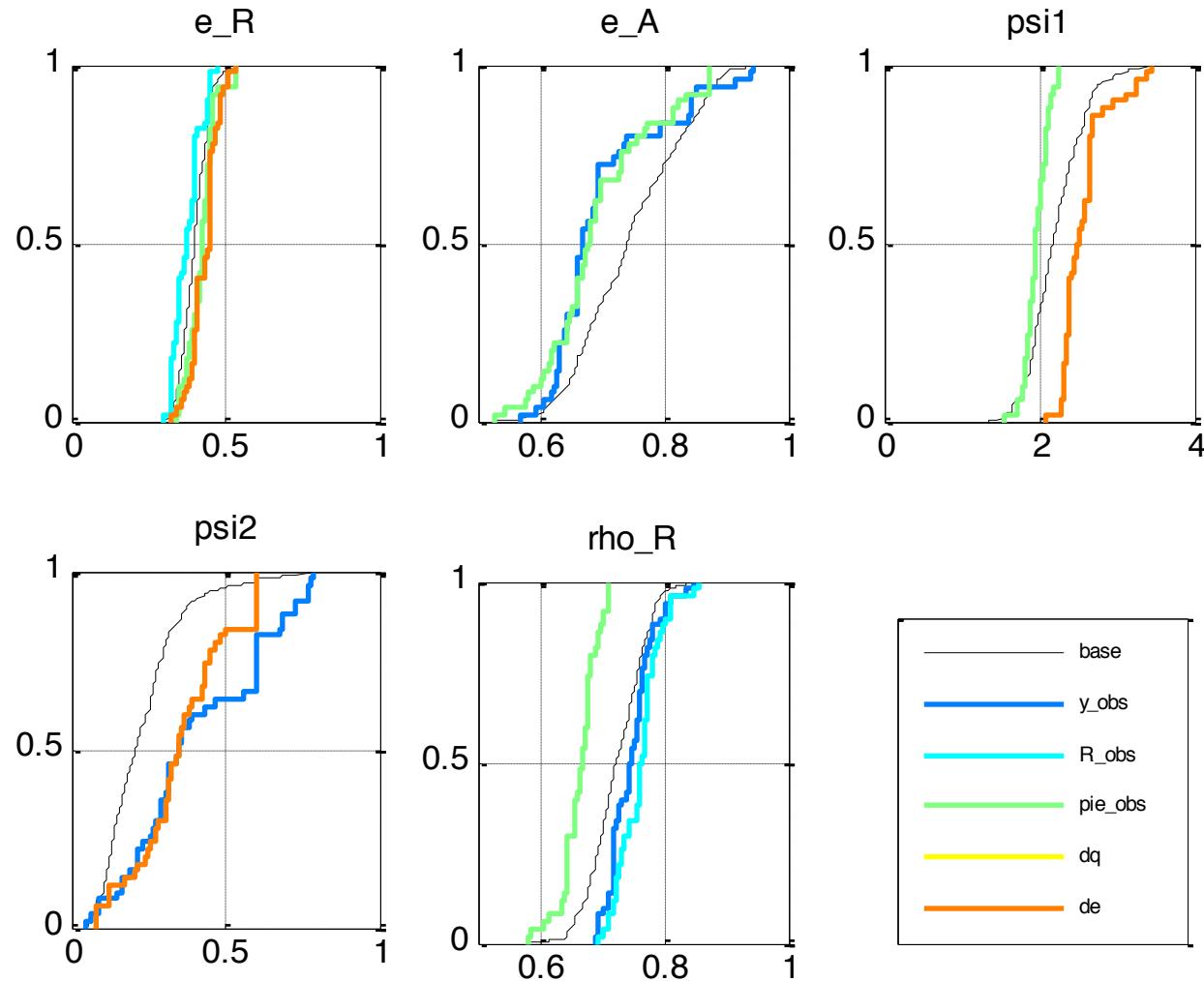
R_{obs}



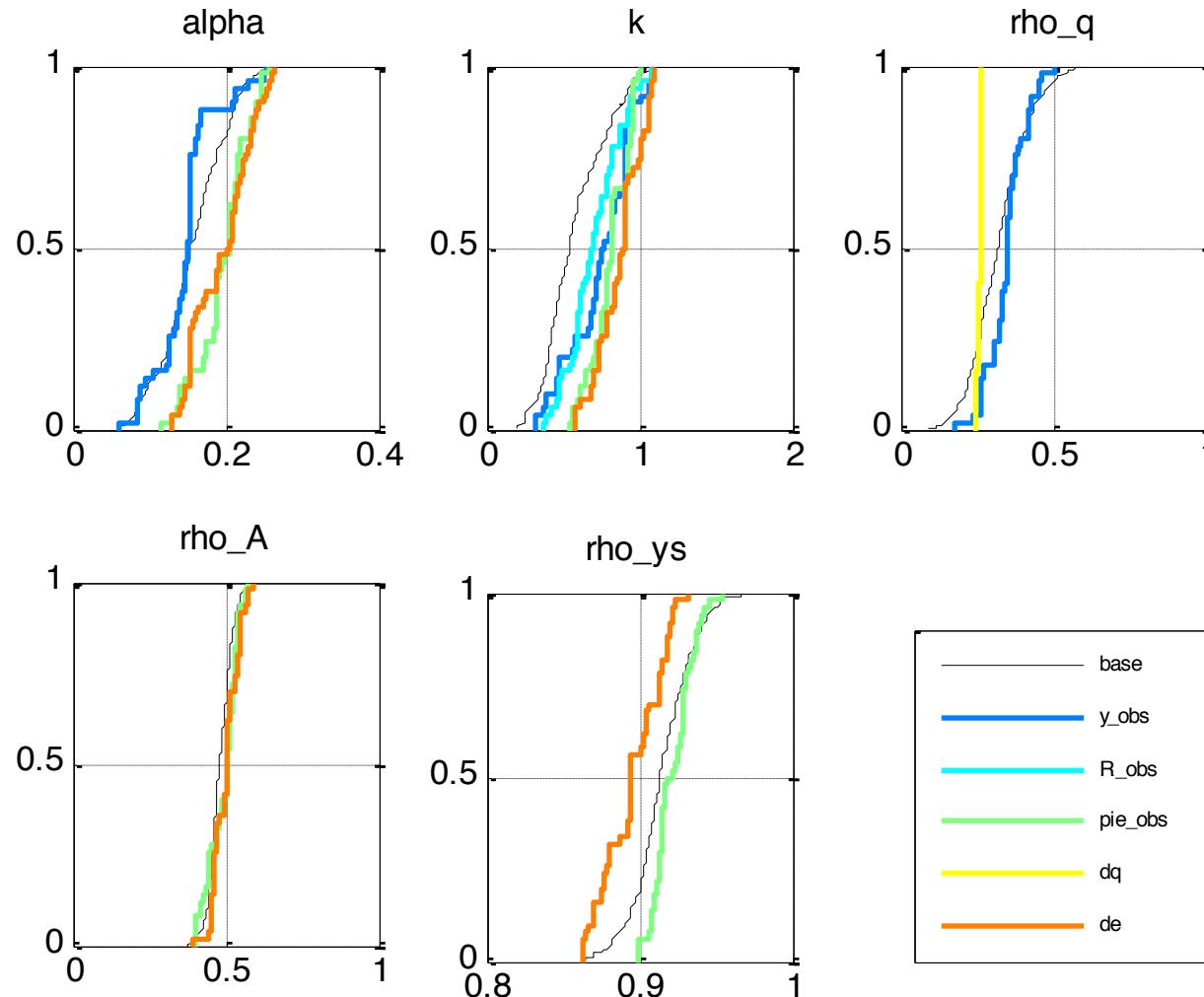
Mapping the fit by observable: prior



Mapping the fit: posterior



Mapping the fit: posterior



Toolbox documentation

option name	default	description
rmse	0	0 = no RMSE analysis 1 = do RMSE analysis
load_rmse	0	0 = make a new RMSE analysis 1 = load previous RMSE analysis
lik_only	0	0 = compute RMSE's for all observed series 1 = compute only likelihood and posterior
var_rmse	varobs	list of observed series to be considered
pfilt_rmse	0.1	filtering threshold for RMSE's: default it to filter the best 10% for each observed series
istart_rmse	1	start computing RMSE's from istart_rmse: use 2 to avoid big initial error
alpha_rmse	0.001	p-value for Smirnov statistics d : plot parameters with p-value < alpha_rmse
alpha2_rmse	0.001	p-value for correlation ρ : plot couples of parameters with p-value < alpha2_rmse

Mapping the reduced form of RE models

Relationship between the reduced form of a rational expectation model and the structural coefficients.

let the reduced form be

$$y_t = Ty_{t-1} + Bu_t,$$

'outputs' Y of our analysis will be the entries in the transition matrix T(X₁,...,X_k) or in the matrix B(X₁,...,X_k).

Lubik Schorfheide (2005)

We analyse the reduced form coefficients describing the relationship between

π_t vs R_{t-1}

We sample the structural coefficients from posterior ranges obtained after estimating the model using data for Canada.

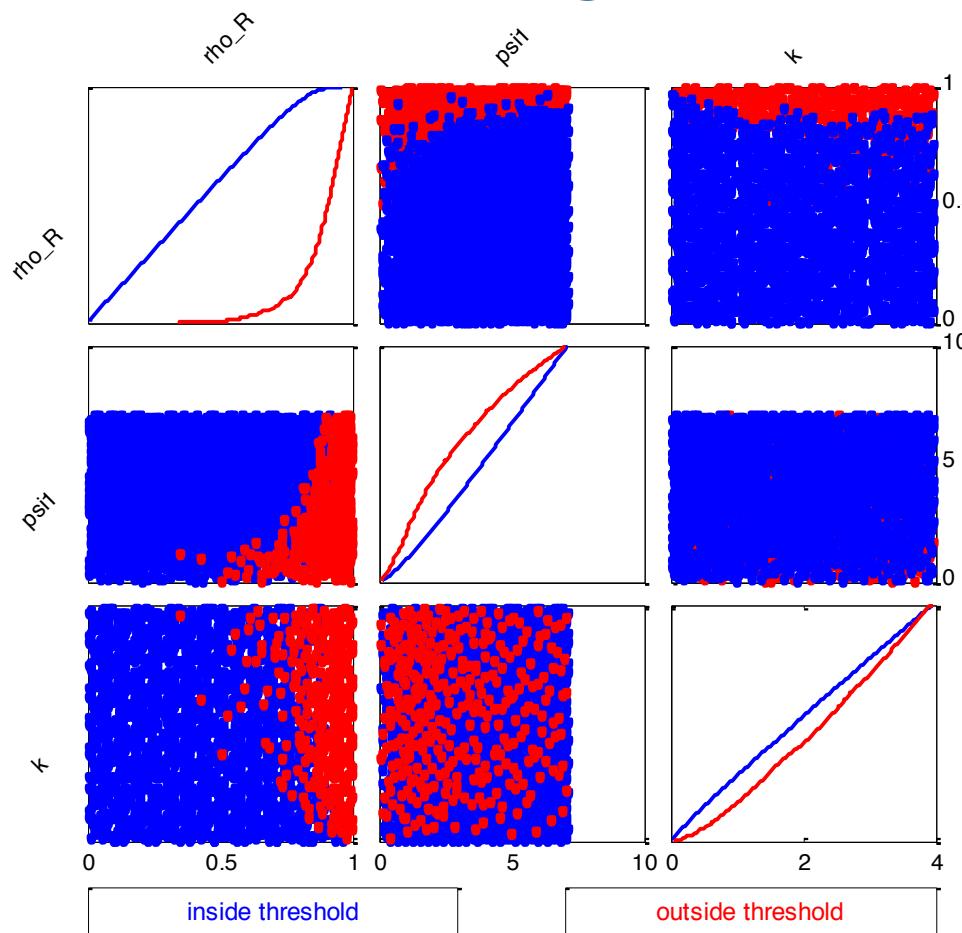
Using Monte Carlo filtering

Store the sample of the state space A,B matrices;

For 1-step ahead irf, perform the MCF sensitivity tests for $B(i,j)$ within/outside the specified ranges.

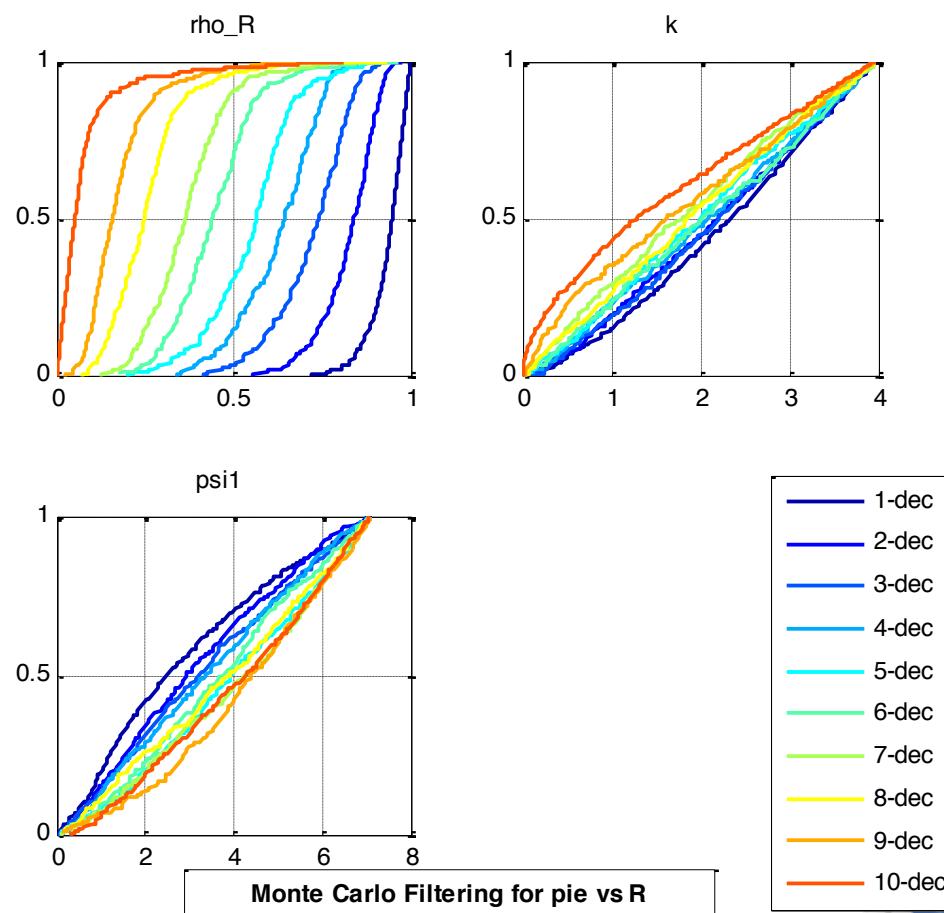
Using Monte Carlo filtering:

$$y = (\eta_t \text{ vs } R_{t-1}) \in [-1, 0]$$



Using Monte Carlo filtering:

$$y = (\pi_t \text{ vs } R_{t-1}) \in [-\infty, \infty]$$



Mapping reduced form solution

option name	default	description
redform	0	0 = don't prepare MC sample of reduced form matrices 1 = prepare MC sample of reduced form matrices
load_redform	0	0 = estimate the mapping of reduced form model 1 = load previously estimated mapping
logtrans_redform	0	0 = use raw entries 1 = use log-transformed entries
threshold_redform	[]	[] = don't filter MC entries of reduced form coefficients [max max] = analyse filtered entries within the range [max max]
ksstat_redform	0.001	critical p-value for Smirnov statistics d when threshold_redform is active
alpha2_redform	0	plot parameters with p-value < ksstat_redform critical p-value for correlation ρ when threshold_redform is active plot couples of parameters with p-value < alpha2_redform
namendo	()	list of endogenous variables
namlagendo	:	jolly character to indicate ALL endogenous
namexo	()	list of lagged endogenous variables: analyse entries [namendo × namlagendo] jolly character to indicate ALL endogenous
	:	list of exogenous variables: analyse entries [namendo × namexo] jolly character to indicate ALL exogenous

RSA: comment

RSA has many 'global' properties:

The whole range of value of input factors is considered, all factors are varied at the same time.

CAVEAT:

RSA classification analyses univariate marginal distributions.

Can miss or underestimate effects of interaction structures.

Note on interactions

$$F(x_1, x_2) = x_1 + x_2 + x_1 * x_2 \quad x_i \sim N(0, 1)$$

- **x1+x2: additive part**
- **x1*x2 : interaction**

RSA: problems

Correlation and interaction structures of the B subset [⇒ also Beck's review, 1987]:

- *Smirnov test is a sufficient test only if H_0 is rejected (i.e. the factor is IMPORTANT);*
- *any covariance structure induced by the classification is not detected by the univariate $d_{m,n}$ statistic.*

e.g. factors combined as products or quotients may compensate

- *bivariate correlation analysis is not revealing, either.*