DSGE models: past, present and future

Fabio Canova

ICREA-UPF, AMeN and CEPR

October 2008

Outline of the talk

- Some history of DSGEs and some general considerations.

- Current problems and a few suggested solutions.

- Future challenges.

References

Canova, F. (2007) What explains the Great Moderation in the United States? A structural analysis, forthcoming, Journal of the European Economic Association.

Canova, F. (2007) How much structure in empirical models, forthcoming, in Mills and Patterson, (eds.) *Palgrave Handbook of Econometrics II*, Palgrave.

Canova, F. (2007) Methods for Applied Macroeconomic Research, Princeton University Press.

Canova, F. (2008) Estimating DSGE models with raw data, manuscript.

Canova, F. and Ferroni, F. (2008) Multiple filtering devices for estimating cyclical DSGE models, manuscript.

Canova, F. and Paustian, M. (2007) Measurement with some theory: using sign restrictions to evaluate business cycles models, manuscript.

Canova, F. and Sala, L. (2005) Back to square one: identification issues in DSGE models, ECB working paper 594, forthcoming, Journal of Monetary Economics.

Most available at www.crei.cat/people/canova

What are DSGE models?

$$E_t[A_\theta x_{t+1} + B_\theta x_t + C_\theta x_{t-1} + D_\theta z_{t+1} + F_\theta z_t] = 0 \quad (1)$$

$$z_{t+1} - G_\theta z_t - e_t = 0 \quad (2)$$

Their stationary (log-linearized) rational expectation solution is:

$$x_t = K_\theta + W_\theta z_{t-1} + J_\theta x_{t-1} + R_\theta e_t \quad (3)$$

$$z_t = G_\theta z_{t-1} + e_t \quad (4)$$

where $K_\theta, W_\theta, J_\theta, R_\theta$ are functions of A_θ, B_θ, $C_\theta, D_\theta, F_\theta$.

Benchmark in academics for:

- Understanding generation and propagation of business cycles.

- Conduct policy analyses.

Popular in central banks/policy circles because:

- Give a coherent story with all general equilibrium interactions.

- Give sharp and easy to communicate optimal policy actions.

- Can use them to forecast (OK when compared to time series models).

- Same language and same tools of academics - no misunderstandings.

History of DSGEs

• 1980's-1990's: Calibrate structural parameters θ, evaluate the fit informally. Do policy analyses with calibrated models.

- Models were simple (driven by very few shocks); no point in estimating them, surely rejected in formal testing.

• 2000's: Estimate structural parameters θ, evaluate the fit formally.

- Models have become more complicated, many shocks, real and nominal frictions.

Big drive to use Bayesian methods. Why?

- Existence of specialized software (Dynare) makes estimation easy (but not always informative).

- Can incorporate external information in the form of a prior - sophisticated interval calibration.

- Classical estimates make sense only if the model is (asymptotically) the DGP of the data up to a set of serially correlated errors.

- Posterior estimates can be obtained without requiring the model to be the DGP; valid in small samples (still, potential interpretation problems if model is misspecified).

- Estimates make sense (not always the case with classical methods).

- Step forward relative to a system of single equations models, estimated by OLS or IV.

- Step forward in terms of understanding the interaction of various economic agents.

- Quality and precision of conditional projections improved.

Open questions

- Parameter identification and use/misuse of priors.

- Mismatch between nature of the model fluctuations and the data.

- Evaluation techniques are invalid under model misspecification.

1. Identification issues

- Crucial DSGE parameters face important identification problems.

- Problems are compounded in large scale models.

- Standard fixup are problematic.

- Bayesian methods are a mixed blessing.

What do we mean by identification?

- Can I transform data information into structural parameters information? Problems can occur in:

• The **solution mapping**, linking the parameters θ to the coefficients of solution, if structural parameters disappear from solution, do not have independent variations, or induce small changes in the solution coefficients.

• The **moment mapping**, linking the coefficients of the solution to the functions of interest (e.g. impulse responses), if the selected functions poorly package the information contained in the solution.

• The **objective function mapping**, linking the functions of interest to the population objective function, if the mapping does not have a unique minimum or if the objective function is "insensitive" to changes in the functions.

• The **data mapping**, linking the population to the sample objective functions, if errors in the identification of shocks or if some variables of the model omitted from the estimated VAR.

Example 1: Observational equivalence, under-identification and weak identification

$$\begin{aligned} y_t &= k_1 + a_1 E_t y_{t+1} + a_2(i_t - E_t \pi_{t+1}) + e_{1t} && (5) \\ \pi_t &= k_2 + a_3 E_t \pi_{t+1} + a_4 y_t + e_{2t} && (6) \\ i_t &= k_3 + a_5 E_t \pi_{t+1} + e_{3t} && (7) \end{aligned}$$

where y_t is the output gap, π_t the inflation rate, i_t the nominal interest rate, e_{1t}, e_{2t}, e_{3t} iid contemporaneously uncorrelated shocks and k_1, k_2, k_3 are constants. The solution is:

$$\begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & a_2 \\ a_4 & 1 & a_2 a_4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{bmatrix} \equiv \mu_\theta + P_\theta e_t \qquad (8)$$

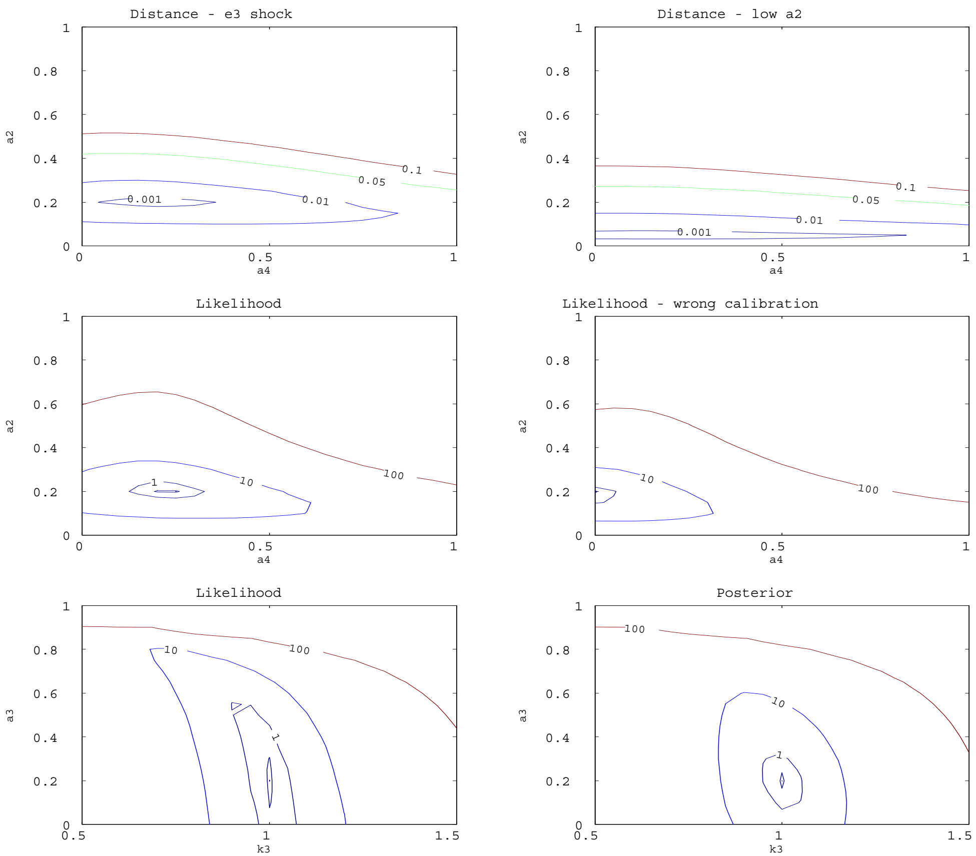
where: $\mu(\theta)$ function of k_i, a_i.

- a_1, a_3, a_5 disappear from the dynamics. Observational equivalence: can't study determinate vs. indeterminate solutions; Sticky vs. non-sticky information.

- Different shocks identify different parameters.

- Different responses carry different information about the parameters.

- ML and distance function (based on impulse responses) have different identification properties. Steady state information matters!



Distance function, Likelihood and Posterior.

• Weak identification of a_4. Partial identification (a ridge) occurs if a_2 is low (see 1.1 and 1.2 boxes)

• Mixed calibration/estimation problematic (true $k_2 = 1$, calibrated $k_2 = 1.2$.) (see 2.2 box)

• Bayesian prior can hide identification problems (compare 3.1 and 3.2 boxes).

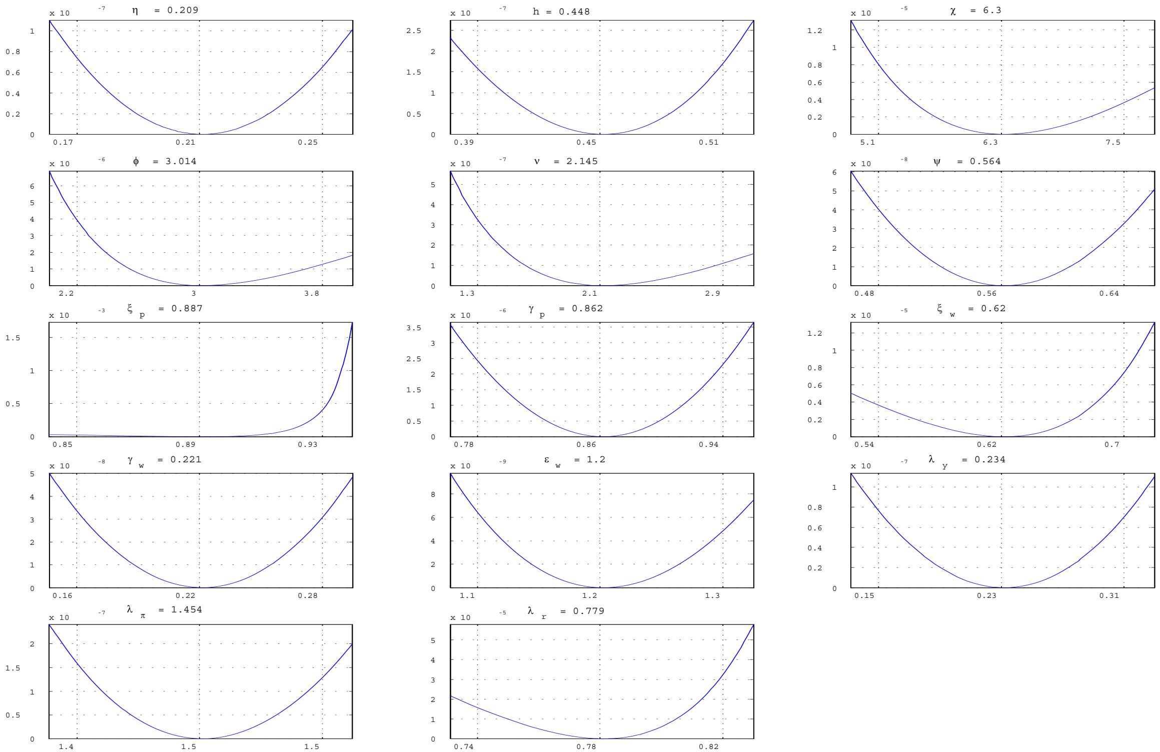
What can one do to solve problems?

- Pathologies occur here because shocks are iid and the model has no internal dynamics. If want to estimate forward looking parameters, add some form of sluggishness or allow serially correlated disturbances.

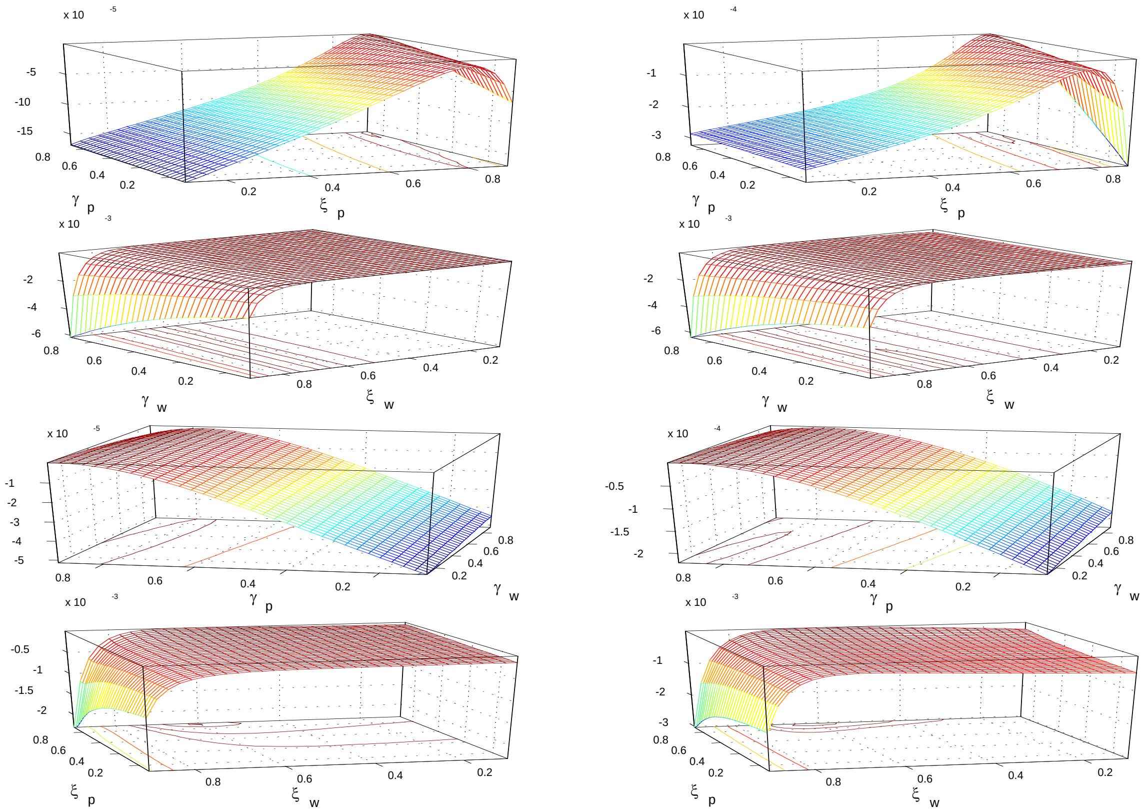
- Always use all possible information (add steady states or the covariance matrix of the shocks information). Limited information procedures likely to have worse identification properties.

- Partial identification problems (ridges) difficult to deal with - need to reparametrize the model.

Example 2: Smets and Wouters model: wrong inference



One dimensional distance functions: monetary policy shocks.



Two dimensional distance functions: monetary policy shocks (first column); all shocks (second column).

What can identification problems lead to?

- Estimate as significant parameters which are missing from the DGP.

- Wrong inference, wrong policy advice, wrong welfare calculations.

	ζ_p	γ_p	ζ_w	γ_w
Baseline	0.887	0.862	0.620	0.221
max	0.867	0.721	0.634	0.279
median	0.857	0.706	0.649	0.349
min	0.888	0.866	0.613	0.206
Case 1	0.001	0.862	0.620	0.221
max	0.004	0.773	0.614	0.132
median	0.000	0.392	0.636	0.259
min	0.021	0.233	0.619	0.214
Case 2	0.001	0.001	0.620	0.221
max	0.000	0.399	0.436	0.175
median	0.134	0.144	0.603	0.143
min	0.001	0.000	0.620	0.221
Case 3	0.001	0.862	0.620	0.001
max	0.000	0.500	0.720	0.260
median	0.019	0.000	0.635	0.052
min	0.006	0.003	0.616	0.001
Case 4	0.887	0.001	0.620	0.800
max	0.000	0.295	0.875	0.050
median	0.889	0.000	0.690	0.434
min	0.887	0.066	0.631	0.658
Case 5	0.887	0.001	0.001	0.221
max	0.174	0.023	0.922	0.062
median	0.886	0.024	0.000	0.144
min	0.887	0.011	0.001	0.049

- With ridges certain parameterizations can never be recovered. Something may turn out to be large when actually is zero.

- Can confuse price and wage stickiness easily.

- Same thing happens with some parameters regulating real rigidities.

2. How to fit cyclical model to trending data?

- Models are typically build to explain only cyclical fluctuations. Why?

- Still too ambitious to account for all types of fluctuations.

- Few known theoretical mechanisms propagating temporary shocks (e.g. R&D as in Comin and Gertler (2006) or Schumpeterian creative destruction as in Canova, Lopez-Salido, Michelacci (2007)).

- Convenient from the computation and the interpretation point of views to assume that the mechanism driving cyclical and non-cyclical fluctuations are distinct and orthogonal.

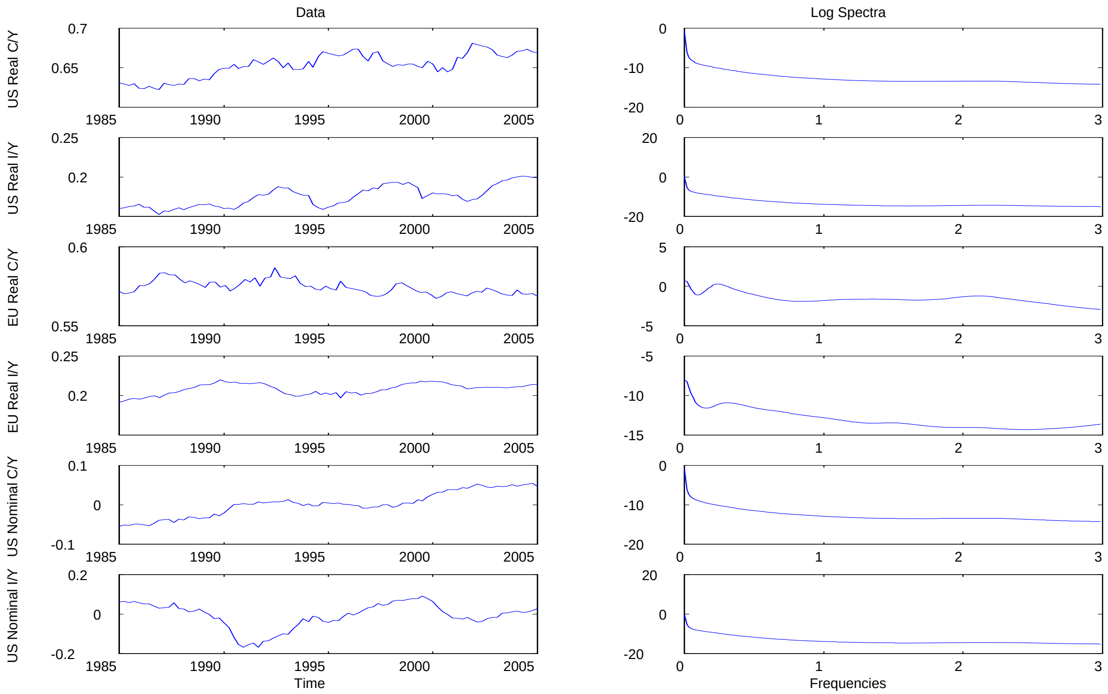
• Mismatch between cyclical DSGE models and the raw data creates headaches.

Two ways to proceed:

• Transform the data taking real or nominal great ratios. Fit the model using transformed data (Cogley (2001), McGrattan (2006)).

- Great ratios are not free of non-cyclical fluctuations.

- They emphasize low frequency fluctuations.



Real and nominal Great ratios in US and EU

• Statistically filter the data. Fit the model to filtered data.

- Which filter should we use? Canova (1998): all filter produce contaminated estimates of fluctuations lasting 8-32 quarters.

- Should we filter each series individually or impose (multivariate) consistency in filtering?

- Should we filter only real or also nominal variables? (in theory non-cyclical shocks may induce cyclical fluctuations in nominal variables, but are all fluctuations in nominal variables cyclical?)

- Many filters are two sided: they change the timing of the shocks.

If estimation results differ, which one shall we trust?

Example 3: A textbook New Keynesian Model (Gali(2008))

- Agents face a labor-leisure choice, there is external habit in consumption.

- Production is carried out with labor.

- Firms face an exogenous probability of price adjustments

- Monetary policy conducted with a conventional Taylor rule.

The log linearized equilibrium conditions are:

$$\begin{aligned}
\lambda_t &= \chi_t - \frac{\sigma_c}{1-h}(y_t - h y_{t-1}) && (9)\\
y_t &= z_t + (1-\alpha) n_t && (10)\\
w_t &= -\lambda_t + \sigma_n n_t && (11)\\
r_t &= \rho_r r_{t-1} + (1-\rho_r)(\rho_\pi \pi_t + \rho_y y_t) + v_t && (12)\\
\lambda_t &= E_t(\lambda_{t+1} + r_t - \pi_{t+1}) && (13)\\
\pi_t &= k_p(w_t + n_t - y_t + \epsilon_t) + \beta E_t \pi_{t+1} && (14)\\
\chi_t &= \rho_\chi \chi_{t-1} + \epsilon_t^\chi && (15)\\
z_t &= \rho_z z_{t-1} + \epsilon_t^z && (16)
\end{aligned}$$

where $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}\frac{1-\alpha}{1-\alpha+\varepsilon\alpha}$, λ is the Lagrangian on the consumer budget constraint, z_t is a technology shock, χ_t a preference shock, v_t is an iid monetary policy shock and ϵ_t an iid markup shock.

The structural parameters:

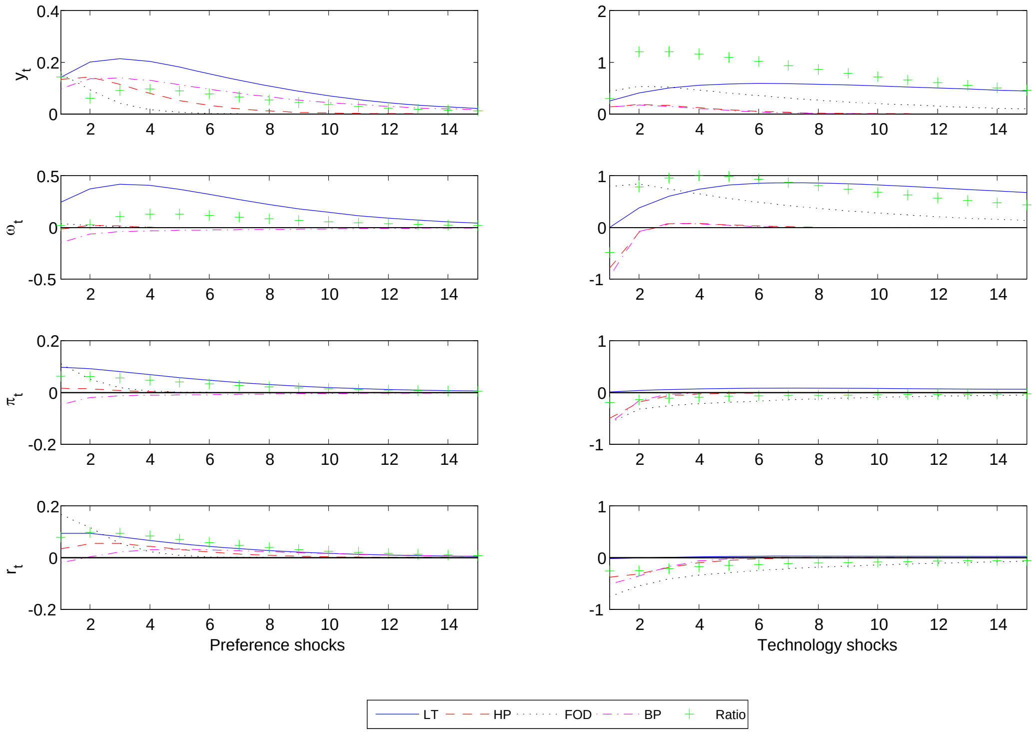
$$\theta = [\sigma_c, \sigma_n, h, 1 - \alpha, \varepsilon, \zeta_p, \rho_r, \rho_\pi, \rho_y, \rho_\chi, \rho_z, \sigma_i, i = 1, \ldots, 4].$$

Observables: (y_t, w_t, π_t, r_t). US data 1980-2006.

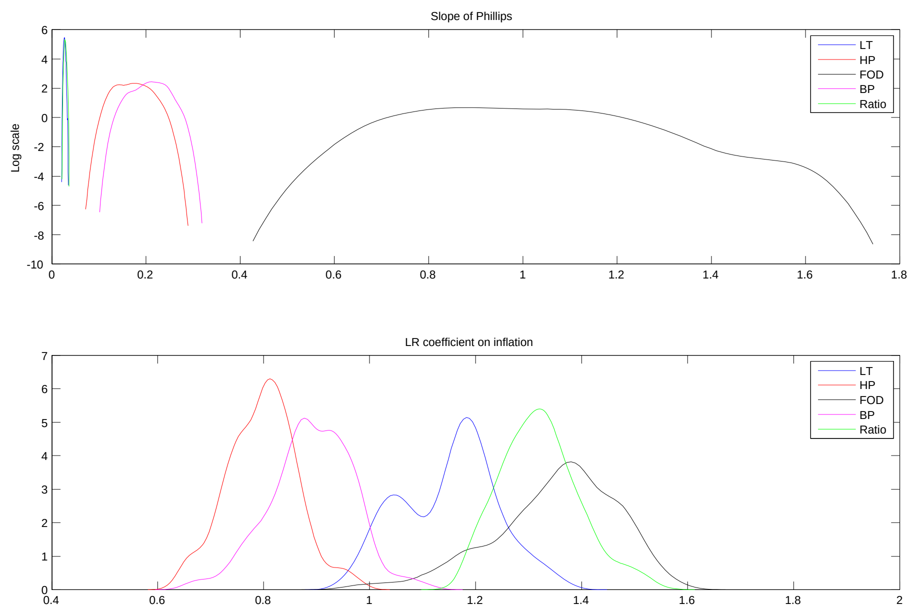
Estimation is conducted with Bayesian MCMC methods.

Filter	LT	HP	FOD	BP	Ratio
	Median (s.d.)	Median (s.d.)	Median (s.d.)	Median(s.d.)	Median(s.d.)
σ_c	2.19 (0.10)	2.25 (0.12)	2.54 (0.16)	2.21 (0.10)	1.69 (0.11)
σ_n	1.79 (0.08)	1.57 (0.10)	1.90 (0.19)	1.78 (0.08)	2.16 (0.10)
h	0.67 (0.01)	0.59 (0.03)	0.44 (0.03)	0.66 (0.02)	0.64 (0.02)
α	0.17 (0.03)	0.12 (0.02)	0.12 (0.03)	0.16 (0.02)	0.13 (0.02)
ϵ	3.90 (0.12)	4.27 (0.14)	2.92 (0.11)	3.72 (0.05)	4.09 (0.12)
ρ_r	0.16 (0.04)	0.52 (0.04)	0.22 (0.06)	0.49 (0.04)	0.22 (0.04)
ρ_π	1.36 (0.08)	1.67 (0.04)	1.74 (0.05)	1.77 (0.08)	1.71 (0.05)
ρ_y	-0.15 (0.02)	0.35 (0.06)	0.13 (0.07)	0.44 (0.05)	-0.02 (0.01)
ζ_p	0.81 (0.01)	0.60 (0.03)	0.33 (0.03)	0.56 (0.03)	0.81 (0.01)
ρ_χ	0.76 (0.02)	0.59 (0.04)	0.29 (0.04)	0.82 (0.03)	0.82 (0.02)
ρ_z	0.96 (0.01)	0.54 (0.05)	0.87 (0.05)	0.46 (0.05)	0.92 (0.01)
σ_χ	0.23 (0.04)	0.37 (0.05)	0.23 (0.04)	0.20 (0.03)	0.95 (0.16)
σ_z	0.12 (0.02)	0.08 (0.01)	0.09 (0.01)	0.09 (0.01)	0.08 (0.01)
σ_{mp}	0.11 (0.01)	0.08 (0.01)	0.12 (0.02)	0.08 (0.01)	0.12 (0.01)
σ_μ	30.54 (1.17)	1.01 (0.40)	0.16 (0.03)	0.63 (0.21)	34.70 (1.04)

Posterior estimates. For LT, HP, FOD and BP real variables detrended, nominal demeaned. For Ratio, real variables are in terms of hours and all variables demeaned.



Impulse responses, different filtering



Phillips curve trade-off and long run inflation coefficient posteriors, different filtering

Why do results depend so much on the preliminary data transformation?

• Filters approximately capture business cycle frequencies in small samples.

i) The LT filter leaves both long and short cycles in the data.

ii) The HP filter leaves high frequencies variability.

iii) The FOD filter emphasizes high frequency noise and downweights the importance of cycles with a business cycle periodicity.

iv) Great ratios leave important low frequency fluctuations in the data

v) Ideal band pass filters in finite samples induce significant approximation errors (see e.g. Canova (2007, ch.3)).

Measurement error is present!

• Filters misspecify the nature and the features of model-based cyclical fluctuations.

i) All filters assume model driven cycle is located at business cycle frequencies and non-cyclical fluctuations are located elsewhere.

ii) The mechanism driving the two types of fluctuations is distinct.

The cyclical component of a DSGE model has power at frequencies other than those corresponding to 8 to 32 quarters. The non-cyclical component induces important fluctuations at business cycle frequencies.

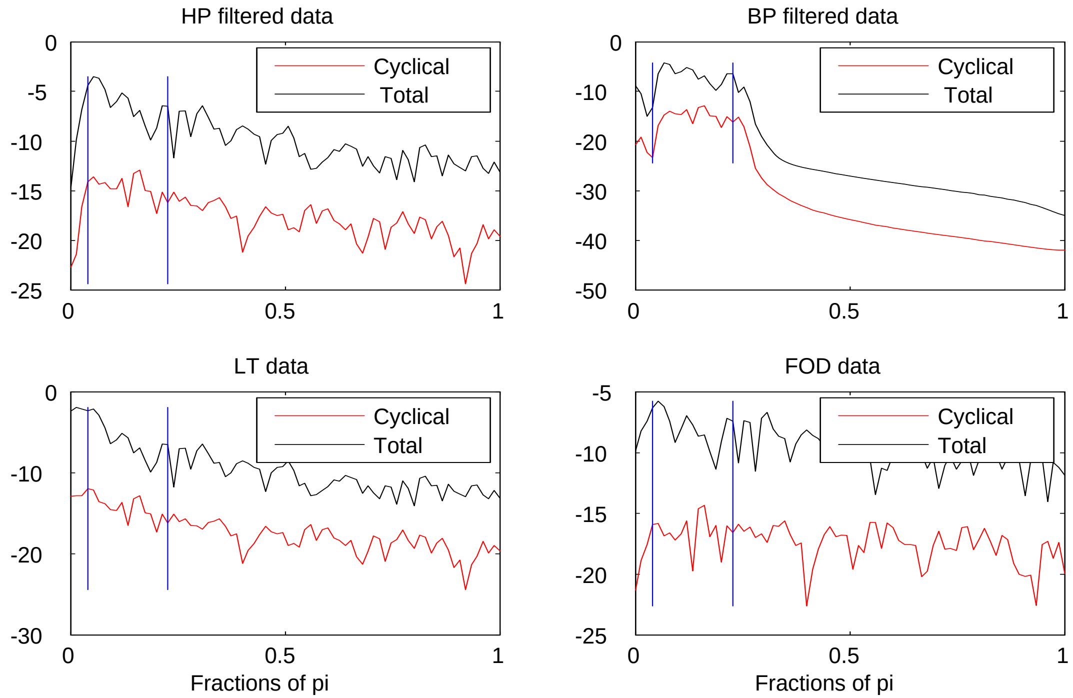
An experiment

- Simulate data from Gali's model, assuming that the preference shock has two components: a stationary and a non-stationary one (Chang, Doh and Schorfheide (2007)).

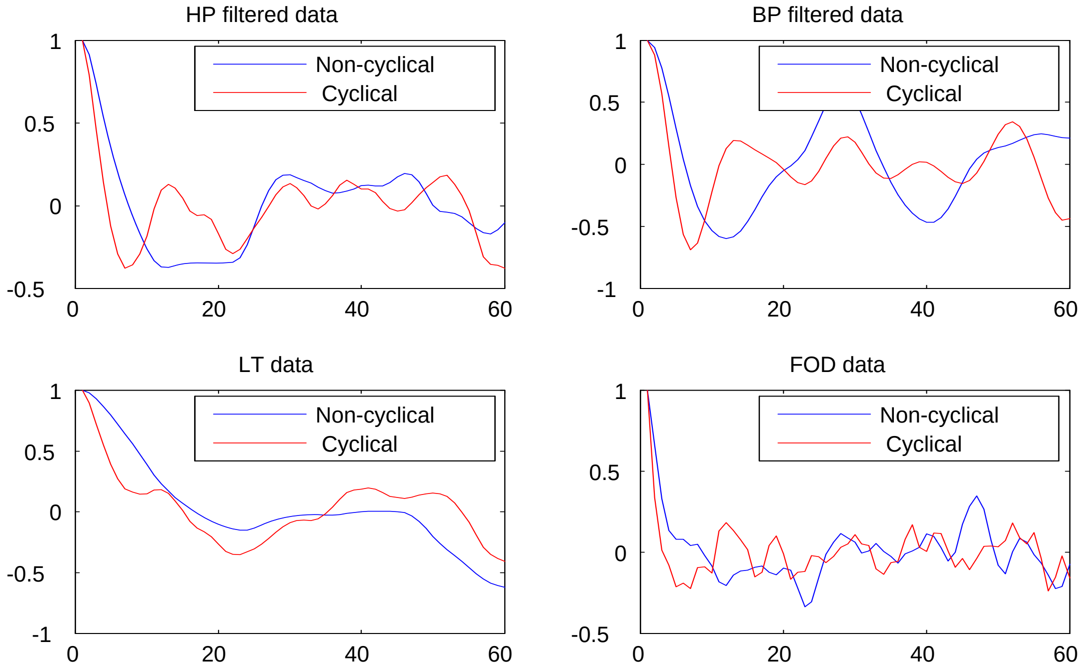
i) the theoretical non-cyclical component corresponds to the fluctuations generated by the non-stationary preference shock.

ii) the theoretical cyclical component corresponds to the fluctuations induced by the other shocks.

- How does the spectrum and autocorrelation function of filtered data looks like?



Model based output log spectra, different filtering. Vertical bars indicate the frequencies where cycles with 8-32 quarters periodicities are located



Autocorrelation function of filtered cyclical and non-cyclical component

- Measurement error present with all filters. .

- The two components have power at all frequencies of the spectrum of output.

- At business cycle frequencies, the theoretical non-cyclical component plays an important role (ACRF of the two components is very similar).

Aguiar and Gopinath (2007): could be an important issue for LDC. Important for all countries!!

Only need that the variability of the shocks driving the two components is of similar magnitude.

What to do? For the first problem: Canova and Ferroni(2008).

Treat all filtered data as contaminated estimate of model quantities.

Let y_t^i be the actual data filtered with method $i = 1, 2, ..., I$ and $y_t^d = [y_t^1, y_t^2, \ldots]$. Assume:

$$y_t^d = \lambda_0 + \lambda_1 y_t(\theta) + u_t \tag{17}$$

where $\lambda_j, j = 0, 1$ are matrices of parameters, measuring bias and correlation between data and model based quantities, u_t are iid measurement errors and θ the structural parameters.

- Factor model setup a-la Boivin and Giannoni (2005).

- Can jointly estimate θ and λ's.

- Can obtain a more precise estimate of the unobserved $y_t(\theta)$ if measurement error is close to iid across methods.

Example 4: What is the role of money for Business cycles?

Use Ireland's (2004) model and US data obtained with 8 different filtering approaches: sample 1961-2008.

Specification	ML	ω_2	ρ_m
Unrestricted	16274	0.44 (0.02)	0.48 (0.02)
$\omega_2 = 0$	16237	0	0.96(0.01)
$\rho_m = 0$	16212	0.43(0.02)	0
$\omega_2 = 0$, $\rho_m = 0$	16220	0	0
Standard estimates		0.03 (0.02)	0.04 (0.03)

Posterior estimates, various specifications.

Second problem: Canova (2008)

Use a flexible link between model and the data

- Model (linearized) solution: cyclical component

$$y_t = RR(\theta)x_{t-1} + SS(\theta)z_t \tag{18}$$

$$x_t = PP(\theta)x_{t-1} + QQ(\theta)z_t \tag{19}$$

$$z_{t+1} = NN(\theta)z_t + \epsilon_{t+1} \tag{20}$$

$PP(\theta), QQ(\theta), RR(\theta), SS(\theta)$ functions of the structural parameters $\theta = (\theta_1, \ldots, \theta_k)$, $x_t = \tilde{x}_t - \bar{x}$; $y_t = \tilde{y}_t - \bar{y}$; and z_t are the disturbances, $\bar{y}, \bar{x}$ are the steady states of $\tilde{y}_t$ and $\tilde{x}_t$

- Non cyclical component

$$y_t^T = y_{t-1}^T + \bar{y}_{t-1} + e_t \qquad e_t \sim \; iid\ (0, \Sigma_e^2) \tag{21}$$
$$\bar{y}_t = \bar{y}_{t-1} + v_t \qquad v_t \sim \; iid\ (0, \Sigma_v^2) \tag{22}$$

$\Sigma_v^2 > 0$ and $\Sigma_e^2 = 0$, y_t^T is a vector of I(2) processes.

$\Sigma_v^2 = 0$, and $\Sigma_e^2 > 0$, y_t^T is a vector of I(1) processes.

$\Sigma_v^2 = \Sigma_e^2 = 0$, y_t^T is deterministic.

$\Sigma_v^2 > 0$ and $\Sigma_e^2 > 0$ and σ_v^2/σ_e^2 is large, y_t^t is "smooth" and nonlinear (as in HP).

- Link between the DSGE model and the observables:

$$y_t^d = c + y_t^T + y_t^m(\theta) + u_t \tag{23}$$

where $y_t^m(\theta) \equiv S[y_t, x_t]'$, S is a selection matrix, $y_t^d \equiv \tilde{y}_t^d - E(\tilde{y}_t^d)$ the log demeaned vector of observables, $c = \bar{y} - E(\tilde{y}_t^d)$, y_t^T is the non-cyclical component, u_t is a iid $(0, \Sigma_u)$ (measurement) noise, $y_t^T, y_t^m(\theta)$ and u_t are mutually orthogonal.

- Jointly estimate structural θ and non-structural parameters.

- Σ_v^2 and Σ_e^2 could be general or restricted e.g. diagonal (so that the non-cyclical component is series specific) or of reduced rank (so that the non-cyclical component is common across series).

Advantages:

• No need to take a stand on the properties of the non-cyclical component and on the choice of filter to tone down its importance - specification errors and biases limited.

• Estimated cyclical component not localized at particular frequencies of the spectrum.

		Small variance			Large variance	
	True	Median	(s.e)	True	Median	(s.e)
σ_c	3.00	3.68	(0.40)	3.00	3.26	(0.29)
σ_n	0.70	0.54	(0.14)	0.70	0.80	(0.13)
h	0.70	0.55	(0.04)	0.70	0.77	(0.04)
α	0.60	0.19	(0.03)	0.60	0.41	(0.04)
ϵ	7.00	6.19	(0.07)	7.00	6.95	(0.09)
ρ_r	0.20	0.16	(0.04)	0.24	0.31	(0.04)
ρ_π	1.30	1.30	(0.04)	1.30	1.25	(0.03)
ρ_y	0.05	0.07	(0.03)	0.05	0.08	(0.10)
ζ_p	0.80	0.78	(0.04)	0.80	0.72	(0.02)
ρ_χ	0.50	0.53	(0.04)	0.50	0.69	(0.05)
ρ_z	0.80	0.71	(0.03)	0.80	0.90	(0.03)
σ_χ	0.011	0.012	(0.0003)	0.011	0.012	(0.0003)
σ_z	0.005	0.006	(0.0001)	0.005	0.007	(0.0001)
σ_{mp}	0.001	0.002	(0.0004)	0.001	0.002	(0.0004)
σ_μ	0.206	0.158	(0.0006)	0.206	0.1273	(0.0004)
σ_χ^{nc}	0.02			0.23		

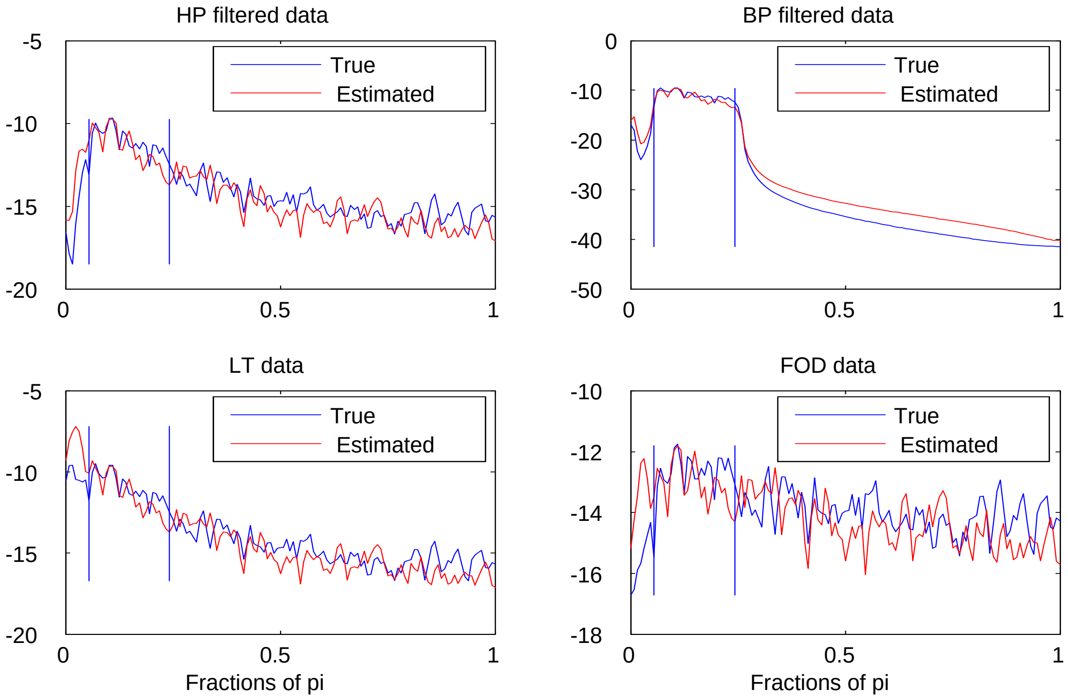
Parameters estimates using flexible specification. σ_χ^{nc} is the standard error of the shock to the non-cyclical component.

Filter		LT	HP	FOD	BP	Flexible
Parameter	True	Median (s.d.)	Median (s.d.)	Median (s.d.)	Median(s.d.)	Median(s.d.)
σ_c	3.00	2.08 (0.11)	2.08 (0.14)	1.89 (0.14)	2.13 (0.12)	3.26(0.40)
σ_n	0.70	1.72 (0.09)	1.36 (0.07)	1.24 (0.06)	1.58 (0.08)	0.54(0.14)
h	0.70	0.67 (0.02)	0.58 (0.03)	0.36 (0.03)	0.66 (0.02)	0.55(0.04)
α	0.60	0.28 (0.03)	0.15 (0.02)	0.14 (0.02)	0.17 (0.02)	0.19(0.03)
ϵ	7.00	3.19 (0.11)	5.13 (0.19)	3.76 (0.18)	3.80 (0.13)	6.19(0.07)
ρ_r	0.20	0.54 (0.03)	0.77 (0.03)	0.72 (0.04)	0.53 (0.03)	0.16(0.04)
ρ_π	1.20	1.69 (0.08)	1.65 (0.06)	1.65 (0.07)	1.63 (0.10)	0.30(0.04)
ρ_y	0.05	-0.14 (0.04)	0.45 (0.04)	0.63 (0.06)	0.40 (0.04)	0.07(0.03)
ζ_p	0.80	0.85 (0.03)	0.91 (0.03)	0.93 (0.03)	0.90 (0.03)	0.78(0.04)
ρ_χ	0.50	1.00 (0.03)	0.96 (0.03)	0.96 (0.03)	0.95 (0.03)	0.53(0.02)
ρ_z	0.80	0.84 (0.03)	0.96 (0.03)	0.97 (0.03)	0.96 (0.03)	0.71(0.03)
σ_χ	1.12	0.11 (0.02)	0.17 (0.02)	0.21 (0.03)	0.14 (0.02)	1.29(0.01)
σ_z	0.51	0.07 (0.01)	0.09 (0.01)	0.09 (0.01)	0.07 (0.01)	0.72(0.02)
σ_{mp}	0.10	0.05 (0.01)	0.05 (0.01)	0.05 (0.01)	0.05 (0.01)	0.22(0.004)
σ_μ	20.60	6.30 (0.50)	16.75 (0.62)	22.75 (0.83)	14.40 (0.58)	15.88(0.06)
σ_χ^{nc}	3.21					

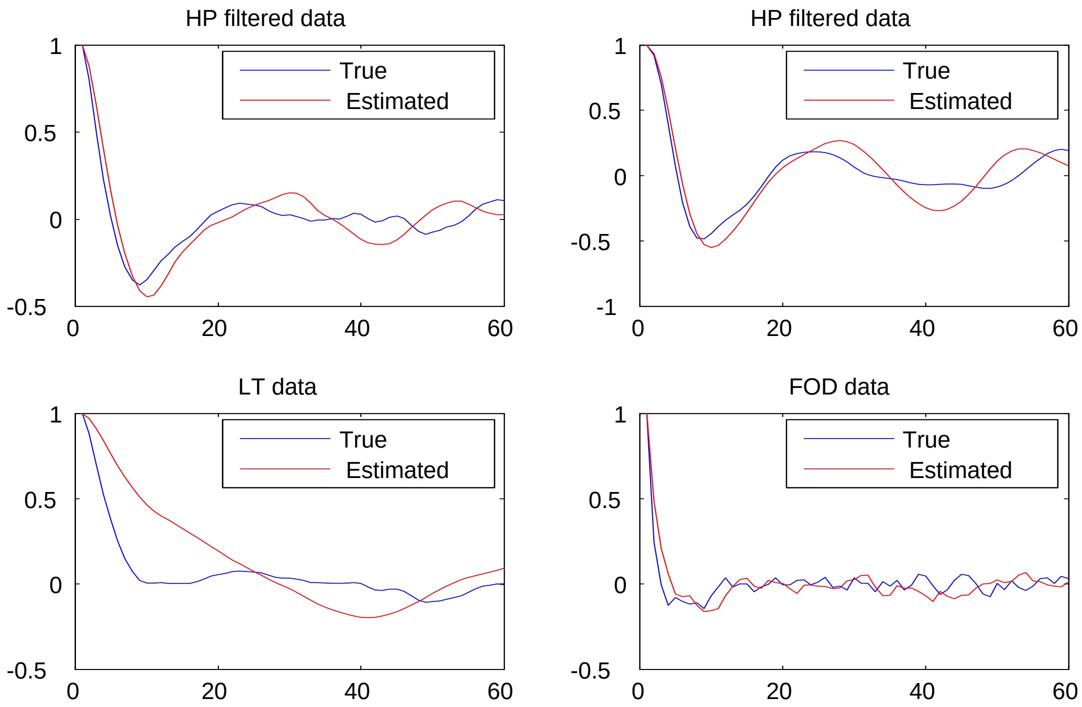
σ_χ^{nc} is the standard deviation of the non-cyclical component. Parameters Estimates using different filters, small variance of non-cyclical shock

Filter		LT	HP	FOD	BP	Flexible
Parameter	True	Median (s.d.)	Median (s.d.)	Median (s.d.)	Median(s.d.)	Median(s.d.)
σ_c	3.00	1.89 (0.07)	1.89 (0.07)	1.87 (0.07)	2.03 (0.09)	3.26 (0.29)
σ_n	0.70	2.13 (0.08)	2.11 (0.08)	2.15 (0.08)	1.90 (0.08)	0.80 (0.13)
h	0.70	0.58 (0.02)	0.60 (0.02)	0.56 (0.02)	0.69 (0.02)	0.77 (0.04)
α	0.60	0.47 (0.02)	0.46 (0.02)	0.49 (0.02)	0.24 (0.03)	0.41 (0.04)
ϵ	7.00	3.85 (0.13)	3.92 (0.13)	3.46 (0.11)	4.16 (0.13)	6.95 (0.09)
ρ_r	0.20	0.68 (0.03)	0.59 (0.03)	0.43 (0.04)	0.50 (0.03)	0.31 (0.04)
ρ_π	1.20	1.14 (0.04)	1.25 (0.04)	1.25 (0.04)	1.23 (0.04)	1.25 (0.03)
ρ_y	0.05	-0.07 (0.00)	-0.01 (0.01)	-0.05 (0.02)	0.23 (0.01)	0.08 (0.10)
ζ_p	0.80	0.81 (0.03)	0.78 (0.03)	0.76 (0.03)	0.89 (0.03)	0.72 (0.02)
ρ_χ	0.50	1.00 (0.03)	1.00 (0.03)	1.00 (0.03)	0.97 (0.03)	0.69 (0.05)
ρ_z	0.80	0.90 (0.03)	0.92 (0.03)	0.91 (0.03)	0.98 (0.03)	0.90 (0.03)
σ_χ	1.12	0.09 (0.01)	0.31 (0.05)	0.61 (0.15)	1.87 (0.14)	1.28 (0.03)
σ_z	0.51	0.61 (0.07)	0.30 (0.04)	0.40 (0.05)	0.10 (0.01)	0.69 (0.01)
σ_{mp}	0.10	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.24 (0.004)
σ_μ	20.60	18.00 (0.74)	18.04 (0.61)	15.89 (0.83)	17.55 (0.57)	12.73 (0.04)
σ_χ^{nc}	23.21					

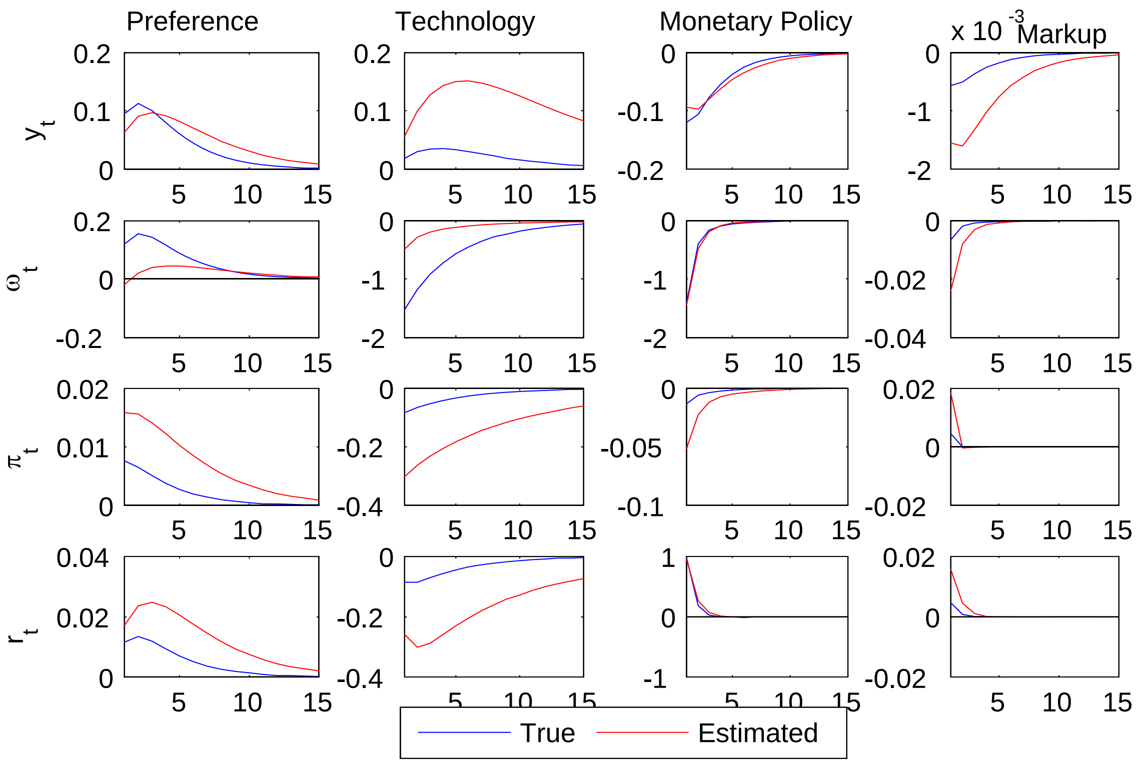
Parameters Estimates using different filters; σ_χ^{nc} is the standard deviation of the non-cyclical component.



Model based cyclical output spectra, true and estimated, different filtering. Vertical bars indicate the frequencies where cycles with 8-32 quarters periodicities are located



Autocorrelation function of filtered cyclical component, true and estimated



Model based IRF, true and estimated.

3. Evaluation problems

- Validation problematic when the model is not the DGP of the data. Standard econometric procedures inapplicable.

- Existing approaches typically use system-wide methods (e.g. LR tests or PO ratios), see Fukac and Pagan (2007).

- Not designed to look at medium run properties of the data, see Kapetanios, Pagan and Scott (2007).

Del Negro and Schorfheide (2004), (2006), Del Negro, et. al. (2006)

• Model can be used as a prior for the data.

• Can verify the quality of model's restrictions by checking how much simulated data must be added to a VAR to improve its fit.

• By-product: can jointly estimate DSGE and VAR parameters.

Problems:

1) Condition on a single model. What if the model is wrong?

2) Still look at the model as a whole.

3) Not much to say about medium run properties of the model.

Canova and Paustian (2007): address problems 1) and 2).

The procedure

- Choose a broad class of structural models which nests submodels through parameter restrictions
  (price and wage stickiness, indexation, habit,...).

- Find implications that are robust to parameter variations in the class.

  – Some implications are robust **across submodels**.

  – Some implications are robust **within a submodel**.

- Use a subset of implications robust across submodels to identify shocks in a SVAR.

- Use implications that are robust within a submodel and different across submodels for evaluation purposes.

- Do this qualitatively and quantitatively using probabilistic criteria.

- Compare submodels (if needed) with the same criteria.

Details

- What are robust restrictions? Signs of the impact response. Magnitude and dynamic restrictions not very robust. Zeros not typically a feature of theory.

- Robust testing: use sign and shape of dynamics of unrestricted variables to shocks.

- Procedure produces a partially identified model: standard statistical criteria problematic (Moon and Schorfheide (2007)).

Why model misspecification not a problem?

- Use robust model based sign restriction.

- Shock identification robust to time series representation of decision rules.

$$
\begin{aligned}
x_{1t} &= A(\theta)x_{1t-1} + B(\theta)e_t \\
x_{2t} &= C(\theta)x_{1t-1} + D(\theta)e_t
\end{aligned}
\tag{24}
$$

$$\begin{bmatrix} I - F_{11}\ell & F_{12}\ell \\ F_{21}\ell & I - F_{22}\ell \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} e_t$$

Representation for y_{2t} (integrating out y_{1t}):

$$(I - F_{22}\ell - F_{21}F_{12}(1 - F_{11}\ell)^{-1}\ell^2)y_{2t} = [G_2 - (F_{21}(1 - F_{11}\ell)^{-1}G_1\ell]e_t \tag{25}$$

ARMA(∞, ∞) but impact effects of e_t have correct sign and magnitude (in population).

Results with experimental data

- Can recognize *qualitative* features of the DGP with high probability.

- Can separate models which are close to each other with high probability.

- Can get a good handle of *quantitative* features of DGP if:

a) Restrictions are abundant (can't be too agnostic).

b) Identified shocks have "relative" large variance.

- Good even in small samples - median response tracks true one well.

Example 5: Hours and technology shocks

- RBC vs. New-Keynesian transmission (Rabanal and Gali (2004) and McGrattan (2004)) after technology shocks.

- Various types of technology shocks (Fisher (2006)), Shumpeterian explanations (Canova, et. al. (2007)).

- What kind of hours dynamics is generated by different types of technology shocks?

- Which type of technology shock drives hours fluctuations most?

Use Christiano, et. al. (2005) and Smets and Wouters (2003) class of models

$$
\begin{aligned}
y_t &= c_y c_t + i_y i_t + g_y e_t^g && (26)\\
c_t &= \frac{h}{1+h} c_{t-1} + \frac{1}{1+h} E_t c_{t+1} - \frac{1-h}{(1+h)\sigma_c}(R_t - E_t \pi_{t+1}) + \frac{1-h}{(1+h)\sigma_c}(e_t^b - E_t e_{t+1}^b) && (27)\\
i_t &= \frac{1}{1+\beta} i_{t-1} + \frac{\beta}{1+\beta} E_t i_{t+1} + \frac{\phi}{1+\beta} q_t - \frac{\beta E_t e_{t+1}^I - e_t^I}{1+\beta} && (28)\\
q_t &= \beta(1-\delta) E_t q_{t+1} - (R_t - \pi_{t+1}) + \beta r^* E_t r_{t+1} && (29)\\
y_t &= \omega(\alpha K_{t-1} + \alpha \psi r_t + (1-\alpha) l_t + e_t^x) && (30)\\
k_t &= (1-\delta) k_{t-1} + \delta i_t && (31)\\
\pi_t &= \frac{\beta}{1+\beta\mu_p} E_t \pi_{t+1} + \frac{\mu_p}{1+\beta\mu_p} \pi_{t-1} + \kappa_p m c_t && (32)\\
w_t &= \frac{\beta}{1+\beta} E_t w_{t+1} + \frac{1}{1+\beta} w_{t-1} + \frac{\beta}{1+\beta} E_t \pi_{t+1} - \frac{1+\beta\mu_w}{1+\beta} \pi_t + \frac{\mu_w}{1+\beta} \pi_{t-1} - \kappa_w \mu_t^W && (33)\\
l_t &= -w_t + (1+\psi) r_t + k_{t-1} && (34)\\
R_t &= \rho_R R_{t-1} + (1-\rho_R)(\gamma_\pi \pi_t + \gamma_y y_t) + e_t^R && (35)
\end{aligned}
$$

	Parameter	Support
σ_c	risk aversion coefficient	[1,6]
h	consumption habit	[0.0,0.8]
σ_l	inverse labor supply elasticity	[0.5,4.0]
ω	fixed cost	[1.0,1.80]
$1/\phi$	adjustment cost parameter	[0.0001,0.002]
δ	capital depreciation rate	[0.015,0.03]
α	capital share	[0.15,0.35]
$1/\psi$	capacity utilization elasticity	[0.1,0.6]
g_y	share of government consumption	[0.10,0.25]
ζ_p	degree of price stickiness	[0.4,0.9]
μ_p	price indexation	[0.2,0.8]
ζ_w	degree of wage stickiness	[0.4,0.9]
μ_w	wage indexation	[0.2,0.8]
ε^w	steady state markup in labor market	[0.1,0.7]
γ_R	lagged interest rate coefficient	[0.2,0.95]
γ_π	inflation coefficient on interest rate rule	[1.1,3.0]
ρ_y	output coefficient on interest rate rule	[0.0,1.0]
$\boldsymbol{\varrho}_i$	persistence of shocks $i = 1, \ldots, 7$	[0,0.9]

Support for the parameters

	TFP	Monetary	Taste	Inv	Markup	L^s	G	
Δy_t	+	+	+	+	+	+	+	
Δc_t	+	+	+	-	+	+	-	
π_t	-	+	+	-	-	-	+	
Δgap_t	+	-	-	?	-	+	-	
Δw_t	+	+	+	-	+	-	?	
Δl_t	-	+	+	+	+	+	+	
R_t	?	-	+	+	?	?	+	
LP_t	+	-	-	-	?	-	-	
i_t	+	+	-	+	+	+	-	
u_t	?	+	+	?	+	?	?	
RR_t	?	-	?	?	?	?	+	

Sign of the impact responses (10^5 draws, 95 % bands).

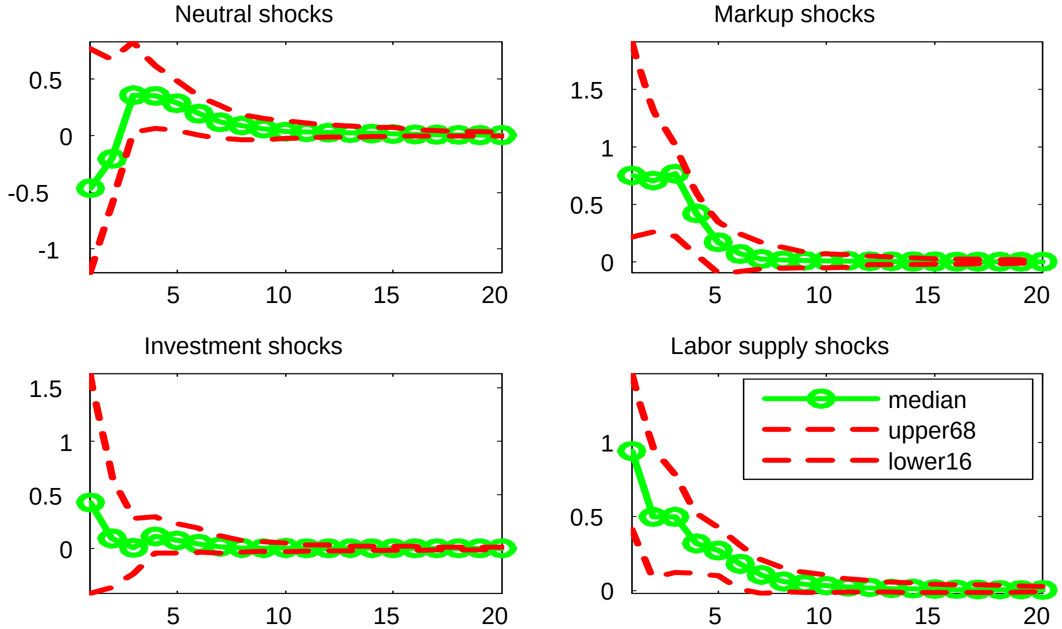
Identification restrictions for technological disturbances:

a) $\pi \downarrow, \Delta y \uparrow$.

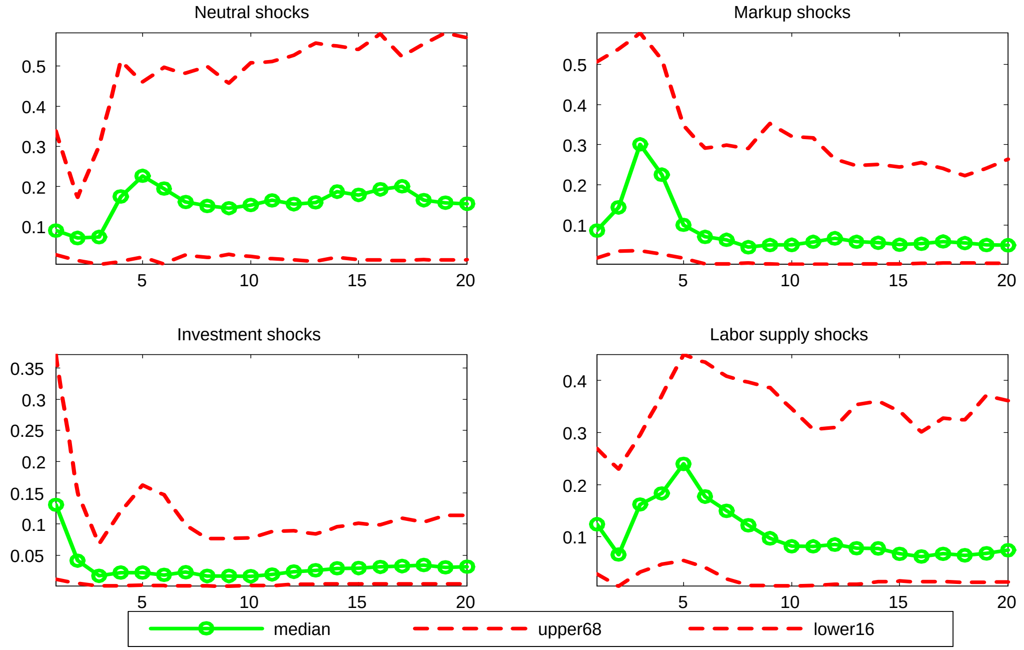
b) $\Delta c \downarrow$ with investment shocks, $\uparrow$ with others.

c) $\Delta gap \uparrow$ with TFP shocks, $\downarrow$ with markup

d) $\Delta w \downarrow$ with L^s and investment shocks, $\uparrow$ markup and TPF shocks.



Responses of hours to technology shocks



Share of hours volatility explained by technology shocks

Example 6: The effects of government spending on consumption

- Models predict that consumption falls after increased government spending: negative wealth effect.

- VAR evidence generally the opposite: Blanchard and Perotti (2002), Fatas and Mihov (2001), Perotti (2007), Pappa (2008).

- Gali et al. (2007): sticky prices and non-Ricardian consumers can produce a rise in consumption following a government spending shock.

- Derive robust restrictions from this class of models. Check if conditioning on the model, consumption increases or falls. (Test of sticky price setup plus presence of rules of thumb consumers).

Sketch of the model by Gali et al (2007)

- fraction $1-\lambda$ of agents are optimizing: $c_t^o = c_{t+1}^o - (r_t - E_t\pi_{t+1})$

- fraction λ of agents are rule of thumb consumers, they consume labor income net of taxes : $c_t^r = \frac{1-\alpha}{\mu\gamma_c}(w_t + n_t^r) - \frac{1}{\gamma_c}t_t^r$

- Calvo price setting: $\pi_t - \mu_p\pi_{t-1} = \kappa(x_t + e_t^u) + \beta(E_t\pi_{t+1} - \mu_p\pi_t)$

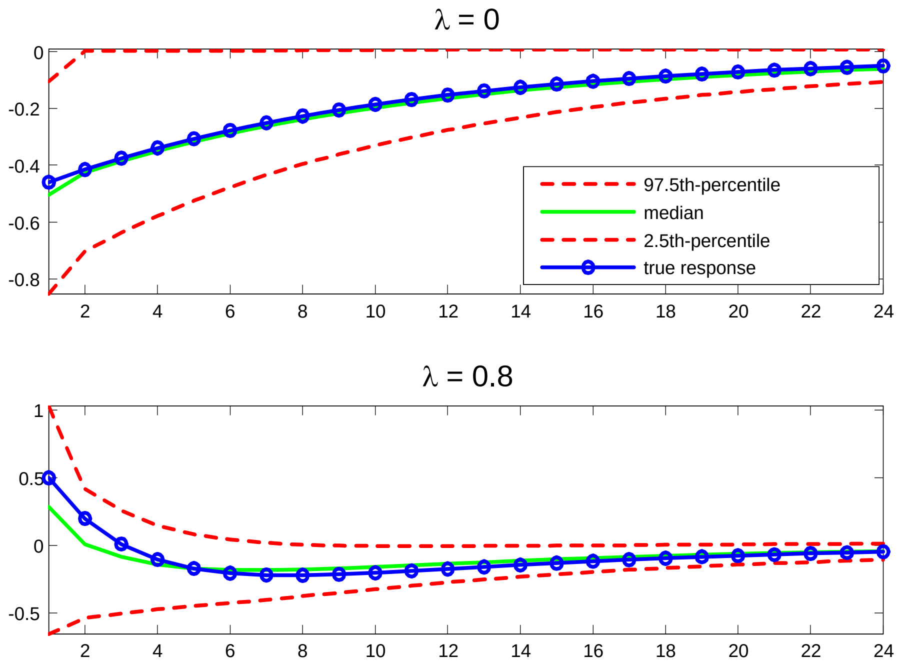
- Sticky prices allow for variable markup.

- Real wages may increases after spending shock despite fall in MPL. Labor income of rule of thumb agents may rise $\rightarrow$ Aggregate consumption may rise

	markup	monetary	**spending**	technology
r	?	?	$+$	-
w	-	-	?	?
π	?	-	**+**	-
y	-	-	$+$	$+$
l	-	-	**+**	-
i	?	?	-	$+$
c	-	-	?	$+$

Signs of impact response (10^5 draws, 95 % bands).

Question: Can the procedure recover the truth?



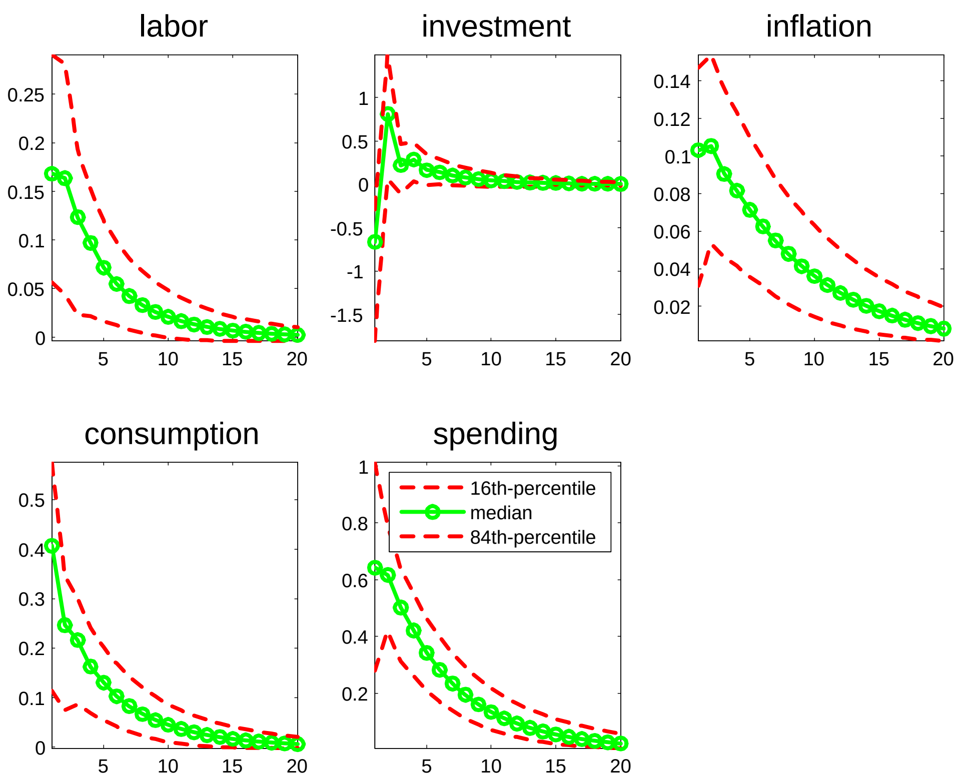
Consumption responses to government spending shock.

U.S. data: 1954:1-2007:1.

- Estimate a 5 variable VAR, ($\{g_t, \pi_t, l_t, c_t, i_t\}$).

- Data is in log differences (except inflation in log levels).

- Identify spending shock ($\{\pi_t > 0, l_t > 0, i_t < 0, g_t > 0\}$) and generic technology shock ($\{\pi_t < 0, l_t < 0, i_t > 0, c_t > 0\}$)



Response to a government spending shock in U.S. data, 1954-2007

Summary

- Don't need to know the exact model, just that it belongs to a class.

- Sign restrictions can be used to evaluate business cycle models.

- Median response is often good estimator of true response.

- Method performs well in small samples. Performance hinges on:

a) shock of interest has relatively large impact on variables (can come from either large variance or impact coefficient in decision rules)

b) impose sufficient number of restrictions (don't be too agnostic)

- Applicable to a number of interesting questions see e.g. Dedola and Neri (2007), Pappa (2007), Gambetti and Pappa (2007).

Future challanges

- For developing countries trend and cycles are the same thing. What is the steady state? How do you proceed? Need to modify solution procedure (find global approximations), estimate models on raw data. Complications.

- Often asymmetric responses are of interest (crisis not the same as a boom). Need nonlinear models to study asymmetries in sign and size of shocks.

- Time varying structures are the norm. Need structural models with Time varying Coefficients. Fernandez Villaverde and Rubio Ramirez (2007) complicated, inference difficult. Canova (2007) easier, but still some interpretation problems.

Thanks for your patience!