Dynare & Bayesian Estimation

Wouter J. Den Haan

University of Amsterdam

August 7, 2010

Overview of the program

- Calculate likelihood, $L(Y^T|\Psi)$
- Calculate posterior, $P(\Psi|Y^T) \propto L(Y^T|\Psi)P(\Psi)$
- Calculate mode
- Calculate preliminary info about posterior
  - using quick and dirty assumption of normality
- Use MCMC to
  - trace the shape of $P(\Psi|Y^T)$
  - calculate things like confidence intervals, t-stats
- Plot graphs

Calculate Likelihood

- Given Ψ get first-order approximation of the model
- Write the system in state-space notation
- Use the Kalman filter to back out

$$\hat{y}_t = y_t - \widehat{\mathrm{E}}\left[y_t | Y^{t-1}, \hat{x}_1\right] \text{ and } \Sigma_{y,t}$$

  - y_t vector with n_y observed values
  - $\widehat{\mathrm{E}}\left[y_t | Y^{t-1}, \hat{x}_1\right]$ prediction according to Kalman filter
  - $\hat{y}_t$: prediction error
  - $\hat{y}_t$: function of *all* the shocks in the model
- Linearity $\Longrightarrow$
  - $\hat{y}_t \sim N(0, \Sigma_{y,t})$
  - likelihood of sequence can be calculated

Calculate posterior & mode

$$P(\Psi|Y^T) \propto L(Y^T|\Psi)P(\Psi)$$

- $P(\Psi|Y^T)$ is a complex function
- But given Ψ, its value can be calculated easily
  - $\Longrightarrow$ value of Ψ that attains the max can be calcualted using an optimization routine

Information about posterior using MCMC

- But we want to calculate objects like

$$E\left[g(\Psi)\right] = \frac{\int g(\Psi)P(\Psi|Y^T)d\Psi}{\int P(\Psi|Y^T)d\Psi}$$

- Examples, mean, standard errors, confidence intervals; these are all integrals

Idea behing MCMC

- generate a sequence for ψ such that
  - its distribution is equal to $P(\Psi|Y^T)$
- But we cannot draw numbers directly from $P(\Psi|Y^T)$
  - implementing MCMC is not trivial

Part I:initialization

```
// dynareestimate.mod

var lc, lk, lz, ly;
varexo e;
parameters beta, rho, alpha, nu, delta;
```

Part II:set values for parameters that are not estimated

```
alpha = 0.36;
rho = 0.95;
beta = 0.99;
nu = 1;
delta = 0.025;
```

- Values will be ignored during estimation
- But maybe you first want to see whether model is sensible by solving it for some parameter values

Part III: model

```
model;

exp(-nu*lc)=beta*(exp(-nu*lc(+1)))
*(exp(lz(+1))*alpha*exp((alpha-1)*lk)+1-delta);

exp(lc)+exp(lk)
=exp(lz+alpha*lk(-1))+(1-delta)*exp(lk(-1));

lz = rho*lz(-1)+e;

end;
```

Part IV (optional): analyze model solution & properties for specified parameter values

```
steady
Stoch_simul(order=1,nocorr,nomoments,IRF=12);
```

- This requires having given numerical values for *all* parameters

Part V: define priors

```
estimated_params;
estderr e, inv_gamma_pdf, 0.007, inf;
end;
```

- more alternatives given below

Part V: initialize estimation

Tell dynare what the observables are

```
varobs lk;
```

Part V: initialize estimation

Give initial values for steady state

```
initval;
lc = -1.02;
lk = -1.61;
lz = 0;
end;
```

Part V: initialize estimation

Steady state must be calculated for many different values of Ψ !!!

- Linearize the system yourself
  - then easy to solve for steady state
- Give the exact solution of steady state as initial values
- Use option to provide a program to calculate the steady state yourself
  - use `modela_steadystate.m` for `modela.mod`

Calculate steady state yourself

```
function [ys,check] = modela_steadystate(ys,exe)
global M_

beta = M_.params(1);
rho = M_.params(2);
alpha = M_.params(3);
nu = M_.params(4);
delta = M_.params(5);
sig = M_.params(6);
check = 0;

z = 1;
k = ((1-beta*(1-delta))/(alpha*beta))^(1/(alpha-1));
c = k^alpha-delta*k;

ys =[ c; k; z ];
```

Part VI: Estimation

Acutal estimation command with some of the possible options

```
estimation(datafile=kdata,mh_nblocks=5,mh_replic=10000,
mh_jscale=3,mh_init_scale=12) lk;
```

- `lk`: (optional) name of the endogenous variables (e.g. if you want to plot Bayesian IRFs)
- `datafile`: contains observables
  - `kdata.mat` or `kdata.m` or `kdata.xls`
- `nobs`: number of observations used (default all)
- `first_obs`: first observation (default is first)

MCMC options

- `mh_replic`: number of observations in each MCMC sequence
- `mh_nblocks`: number of MCMC sequences
- `mh_jscale`: variance of the jumps in Ψ in MCMC chain
  - a higher value of `mh_jscale` $\Longrightarrow$ bigger steps through the domain of Ψ & lower acceptance ratio
  - acceptance ratio should be between 0.2 and 0.4
- `mh_init_scale:` variance of initial draw
  - important to make sure that the different MCMC sequences start in different points

Part VII: Using estimated model

Acutal estimation command with some of the possible options

shock_decomposition;

- `plots graphs with the observables and the part explained by which shock.`

Priors - Format

```
estimated_params;
parameter name, prior_shape, prior_mean,
prior_standarddeviation
      [,prior 3rd par. value, prior 4th par. value];
end
```

name	distribution & parameters	range
normal_pdf	$N(\mu,\sigma)$	$\mathbb{R}$
gamma_pdf	$G_2(\mu,\sigma,p_3)$	$[p_3,+\infty)$
beta_pdf	$B(\mu,\sigma,p_3,p_4)$	$[p_3,p_4]$
inv_gamma_pdf	$IG_1(\mu,\sigma)$	$\mathbb{R}^+$
uniform_pdf	$U(p_3,p_4)$	$[p_3,p_4]$

Priors - Examples

alpha's prior is Normal with mean `mu` and standard deviation `sigma`:

```
alpha , normal_pdf,mu,sigma;
```

alpha's prior is uniform over `[p3,p4]`:

```
alpha , uniform_pdf, , ,p3,p4;
```

Note the two spaces betweeen the commas

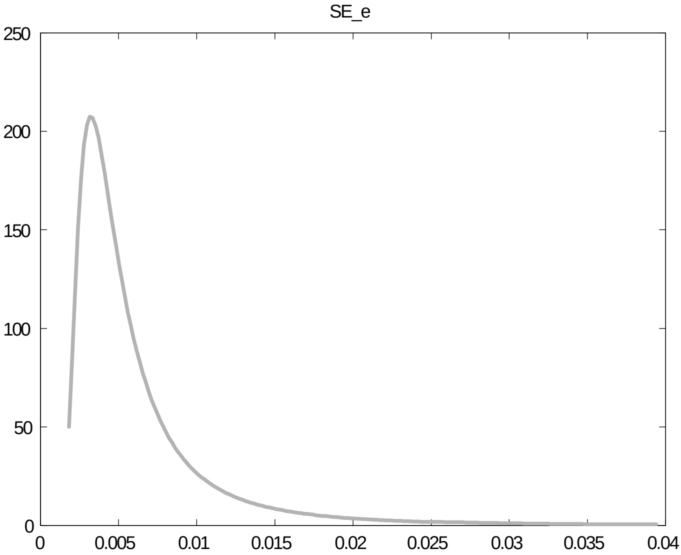
What to watch while Dynare runs

Plots here from two examples

1. As good as it gets: estimate only 1 parameter
2. Estimate all parameters

- Both cases 5,000 observations
- No misspecification of the model
  - i.e., artificial data

What to watch while Dynare runs



What to watch while Dynare runs

- When Dynare gets to the MCMC part a window opens telling you
    - in which MCMC run you are
    - which fraction has been completed
    - **most importantly** the acceptance rate
- The acceptance rate should be between 0.2 & 0.4
    - a relatively low acceptance rate makes it more likely that the MCMC travels through the whole domain of Ψ
    - acceptance rate too high $\Longrightarrow$ increase `mh_jscale`

Tables

- RESULTS FROM POSTERIOR MAXIMIZATION
  - generated before MCMC part
  - most important is the mode, the other stuff is based on normality assumptions which are typically not valid
- ESTIMATION RESULTS (based on MCMC)

Graphs

- Prior
- MCMC diagnostics (see below)
- Prior & posterior densities
- Shocks implied at the mode
- Observable and corresponding implied value

Brooks & Gelman 1989 statistics

- MCMC: should give sequence with draws from a fixed distribution
- Thus, distribution or moments should be similar
  - for different parts of the same sequence
  - across sequences (if you have more than one)

Brooks & Gelman 1989 statistics

- Ψ_{ij} the i^{th} draw (out of I) in the j^{th} sequence (out of J).

$$
\begin{aligned}
B &= \frac{I}{J-1}\sum_{j=1}^{J}\left(\bar{\Psi}_{.j}-\bar{\Psi}_{..}\right)^2 \\
W &= \frac{1}{J}\sum_{j=1}^{J}\frac{1}{I-1}\sum_{i=1}^{I}\left(\bar{\Psi}_{ij}-\bar{\Psi}_{.j}\right)^2 \quad \text{(red line)} \\
\widehat{VAR} &= \frac{I-1}{I}W+\frac{1}{I}B \quad \text{(blue line)}
\end{aligned}
$$

Between variance

$$\bar{B} = \frac{1}{J-1}\sum_{j=1}^{J}\left(\bar{\Psi}_{\cdot j} - \bar{\Psi}_{\cdot\cdot}\right)^2$$

- $\bar{B}$ is an estimate of the variance of the mean $(\sigma^2/I) \Longrightarrow B = \bar{B}I$ is an estimate of the variance

Within variance

$$W = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{I-1}\sum_{i=1}^{I}\left(\bar{\Psi}_{ij} - \bar{\Psi}_{.j}\right)^2$$

- W is an estimate of the variance (averaged across streams)

Weighted combination

Both B and W are estimates of σ^2 and so is

$$\widehat{VAR} = \frac{I-1}{I}W + \frac{1}{I}B$$

Interpretations:

- If B and W converge then $\widehat{VAR}$ and W also converge
- If the variance of the mean across streams goes to zero as m increases, then $B/I = \bar{B}$ converges to zero

Brooks & Gelman 1989 statistics

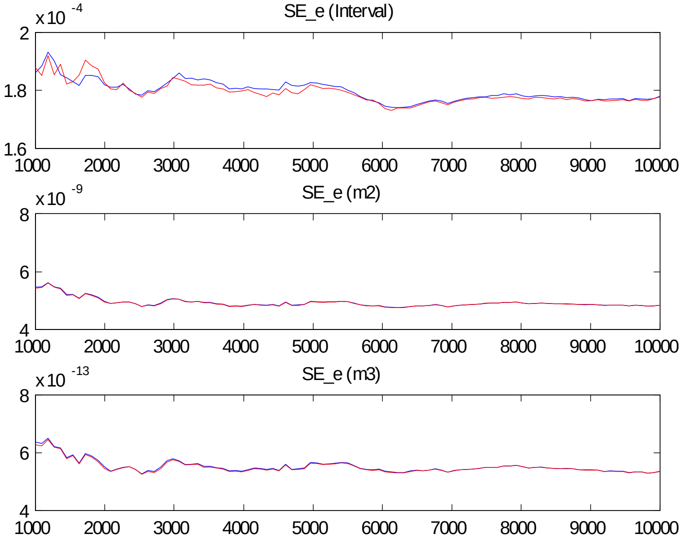
What do we want them to do?

- Settle down (variation within sequences disappears)
- Red and blue line converge (consistency across sequences)

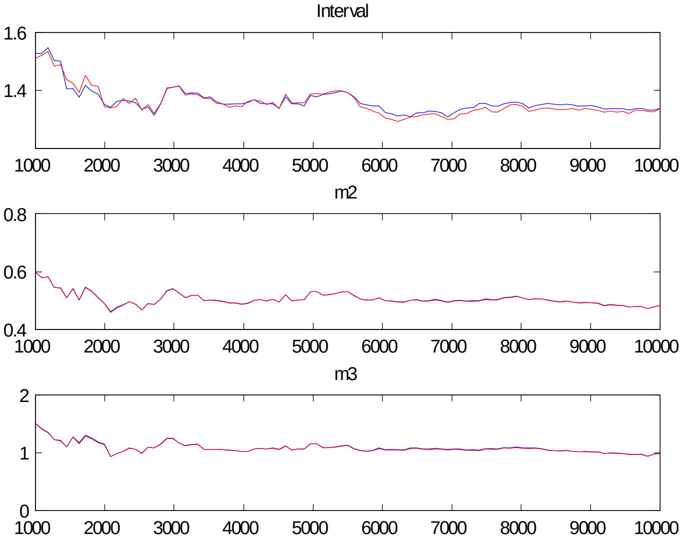
Output for first example

- Estimate only 1 parameter, standard deviation innovation
- correctly specified (neoclassical growth) model
- 5,000 observations

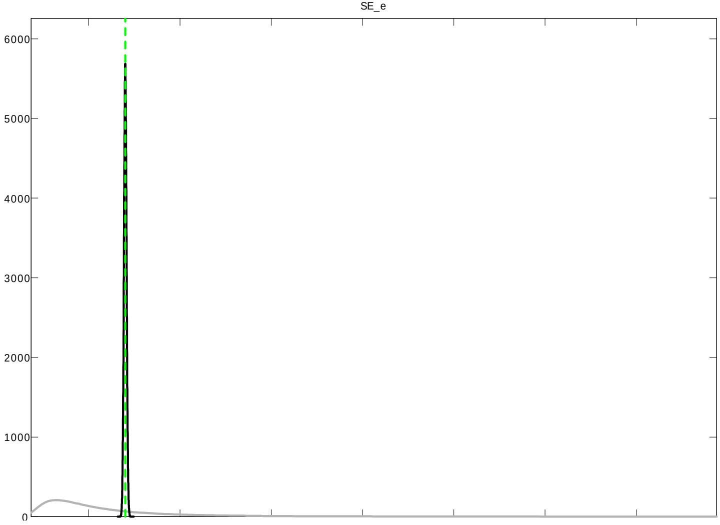
Univariate MCMC diagnostics I



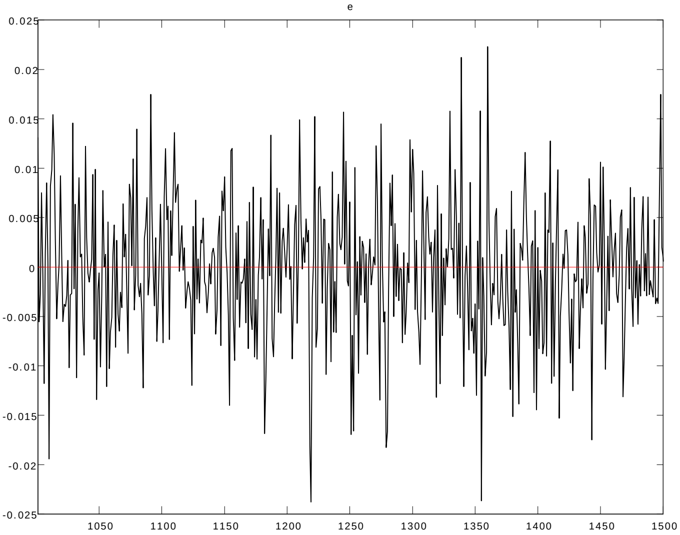
Multivariate MCMC diagnostics II



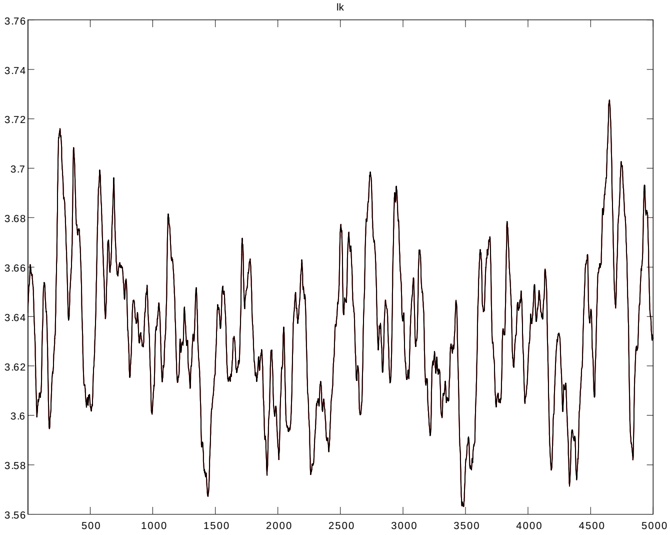
Posterior densities



Shocks


e
0.025
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
-0.02
-0.025
1050
1100
1150
1200
1250
1300
1350
1400
1450
1500

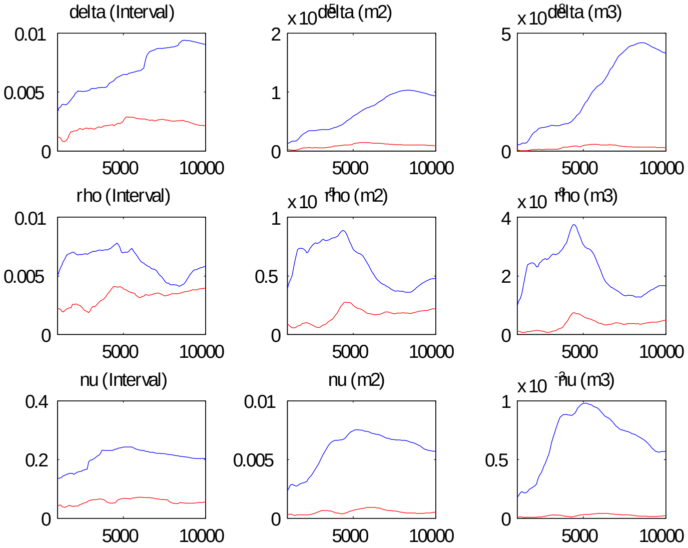
Observables and implied values



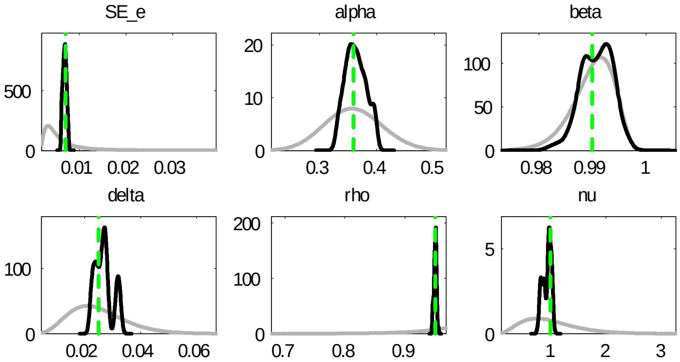
Output for second example

- Estimate all parameters
- correctly specified (neoclassical growth) model
- 5,000 observations

Second example - MCMC



Second example - Posteriors



How to make the MCMC pics look better?

Problem:

- Parameters not well identified, possibly because the dynamics of the model are too simple; capital is not much more than a scaled up version of productivity
- More data doesn't seem to help

Are dynamics caused by model or shocks?

- What explains the data, the shocks or the model?
- How much propagation does the model really have?
- Two examples:
    - Standard RBC
    - Christiano, Motto, Rostagno

Propagation in standard RBC

Policy rule in DSGE model:

$$x_{t+1} = a_0 + A_1 x_t + A_2 shocks_t$$

- Propagation, i.e., economic theory is all in A_1
- Exogenous stuff is in $A_2 shocks_t$
- Economics important $\Longrightarrow$

$$\tilde{x}_{t+1} = \tilde{a}_0 + \tilde{A}_2 shocks_t$$

  should give a bad fit

  - no matter what values of $\tilde{a}_0$ and $\tilde{a}_1$ used

Check importance of economics in your model

- Let

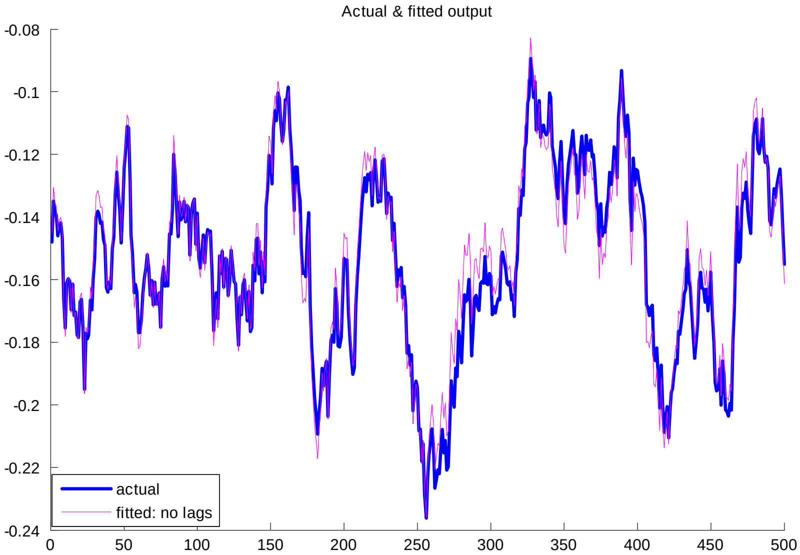
$$\arg\min_{\tilde{a}_0,\tilde{a}_1} \sum_{t=2}^{T} \left(x_{t+1} - \tilde{a}_0 - \tilde{A}_2 shocks_t\right)^2$$

- plot $\{x_{t+1}, \tilde{x}_{t+1}\}$

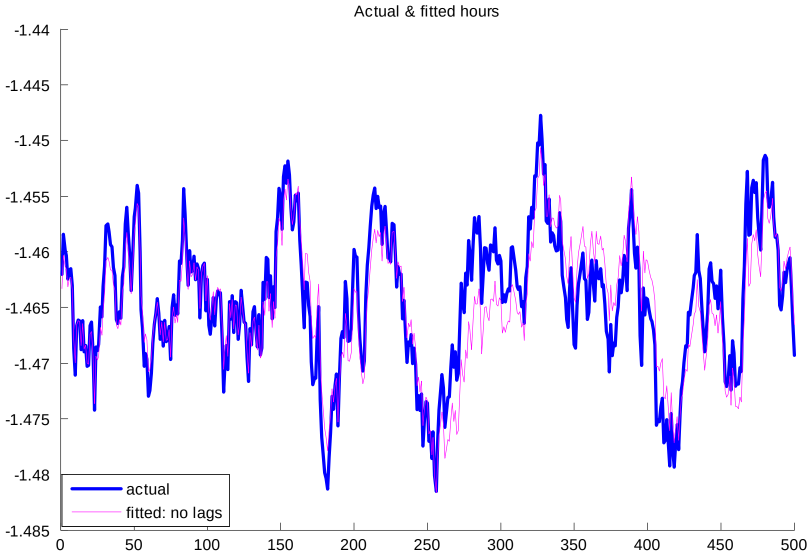
Example

- Propagation in standard growth model
- Why would the endogenous variables not follow driving process 1 for 1?

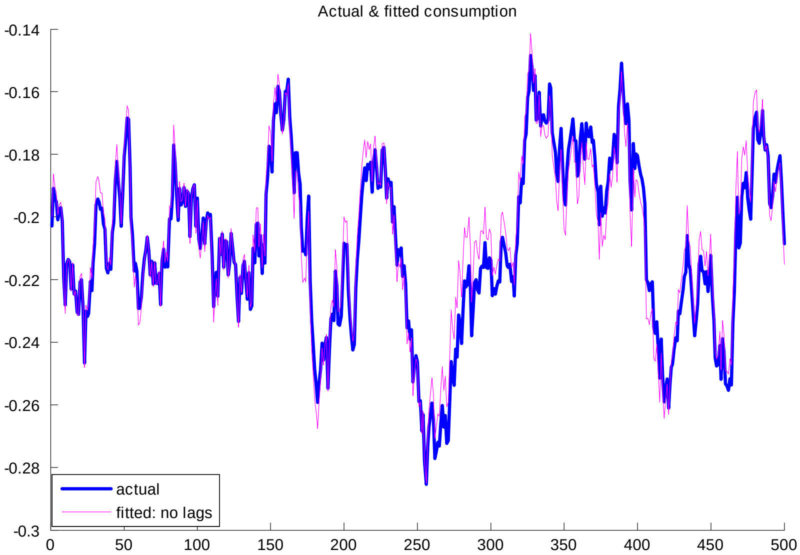
Output & current productivity shock



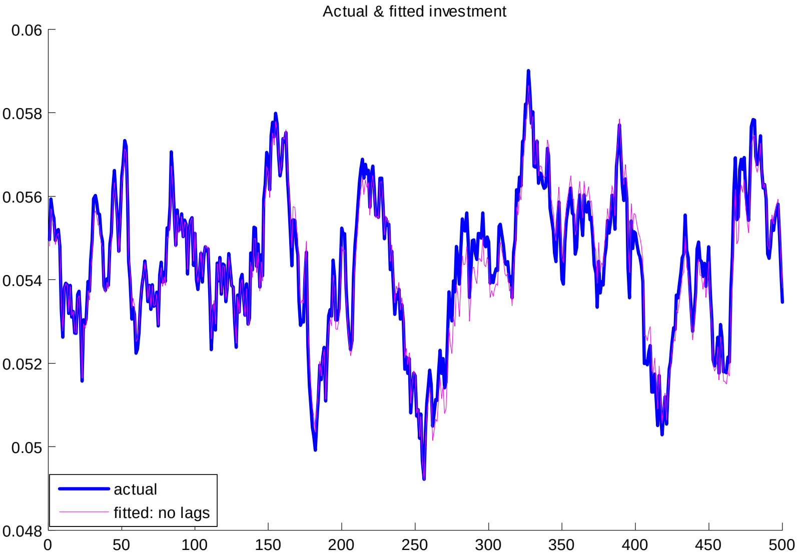
Hours & current productivity shock



Consumption & current productivity shock

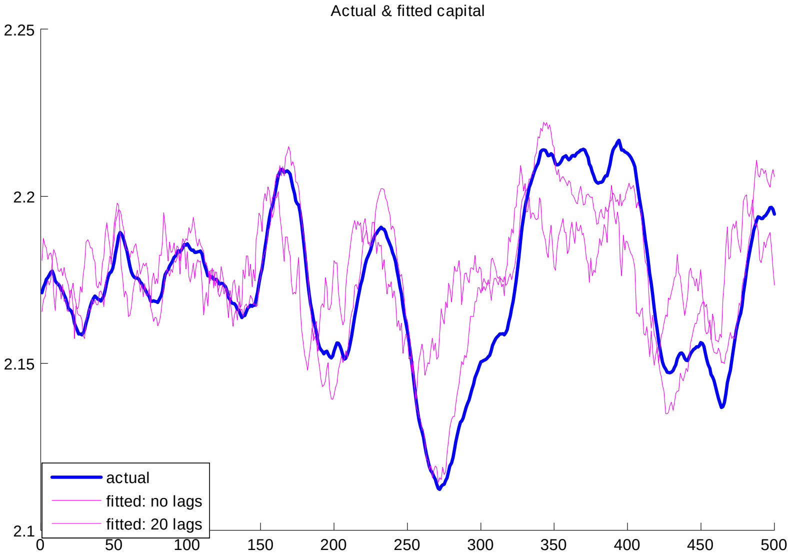


Investment & current productivity shock

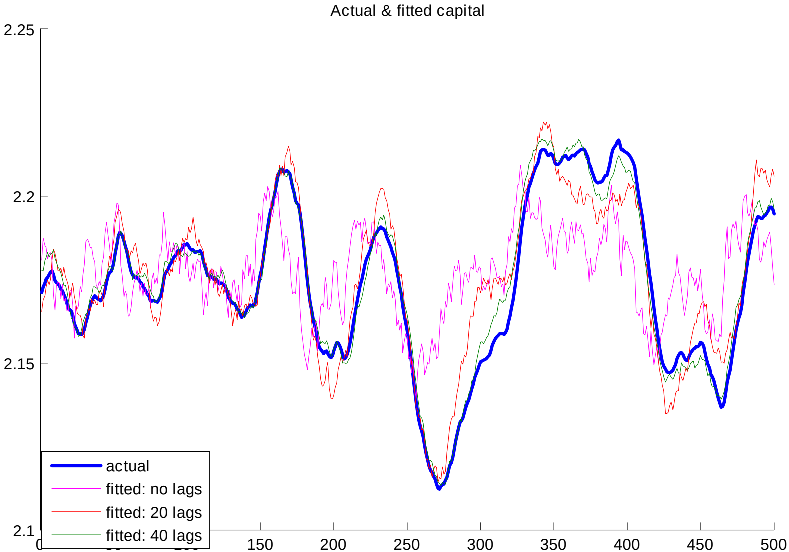


Capital & current productivity shock

Adding 20 lagged values of the shock



Adding 40 lagged values of shock



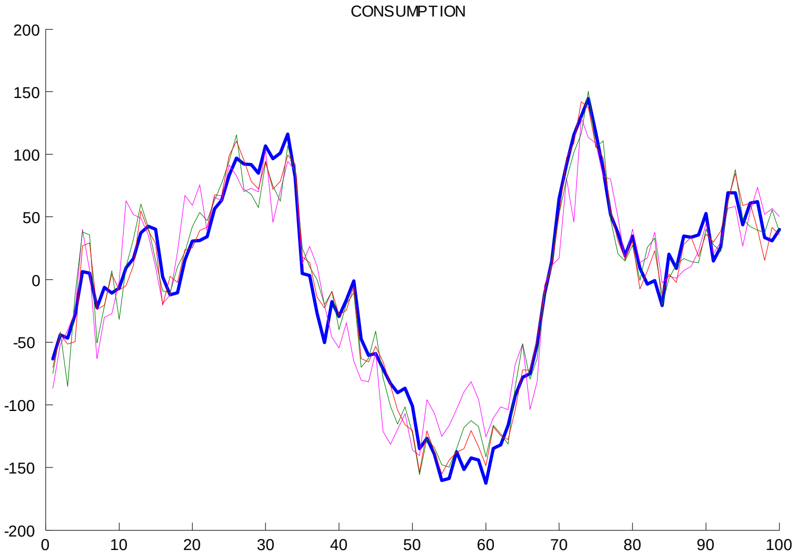
R2

	Only current	+20 lags	+40 lags
Output	0.958	0.993	0.999
Hours	**0.825**	0.971	0.994
Consumption	0.947	0.991	0.998
Investment	0.966	0.994	0.999
Capital	**0.288**	**0.880**	0.976

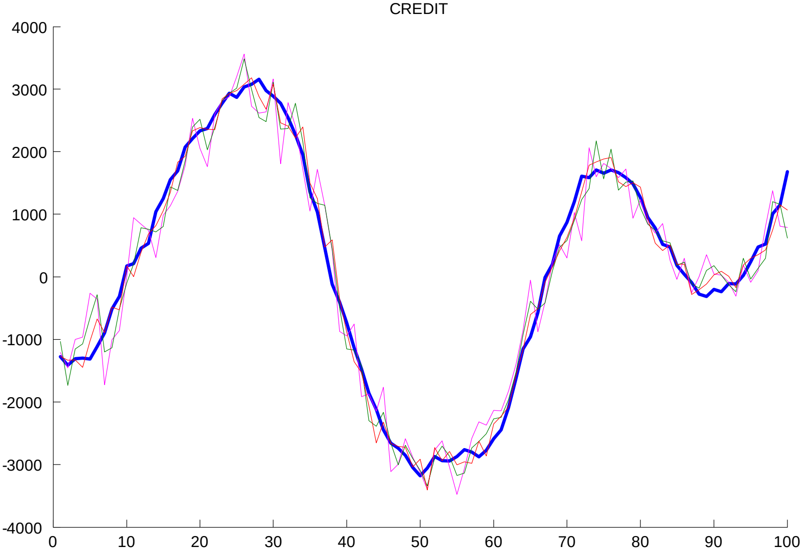
Second example

- Christiano, Motto, Rostagno:
  - "Financial Factors in Economic Fluctuations"
- Quite complex model to model interaction between financial intermediation and real activity

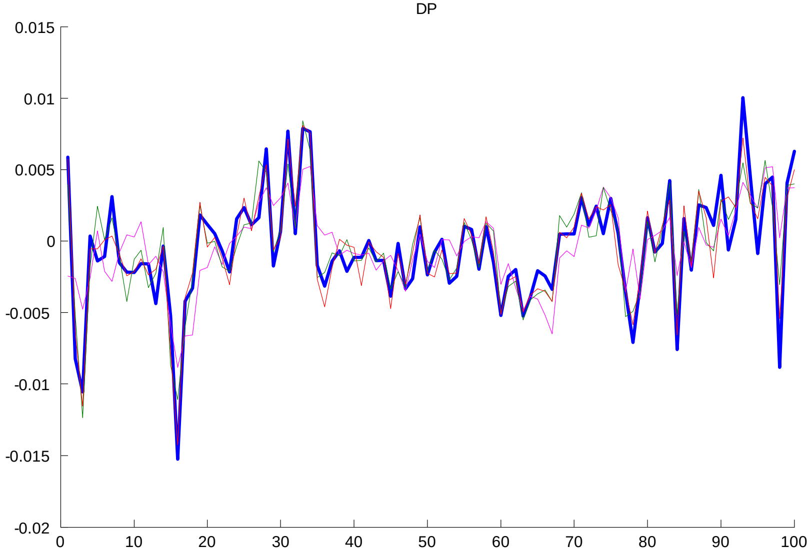
Using current shocks & 1 & 2 lags



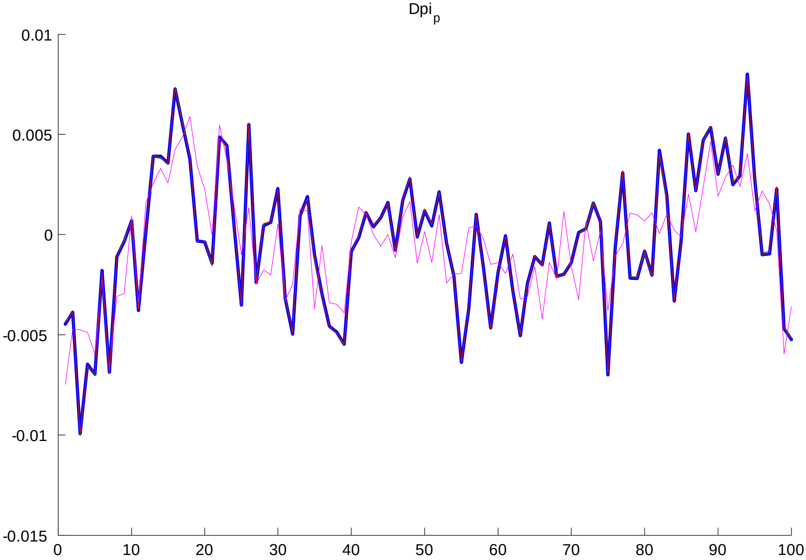
Using current shocks & 1 & 2 lags



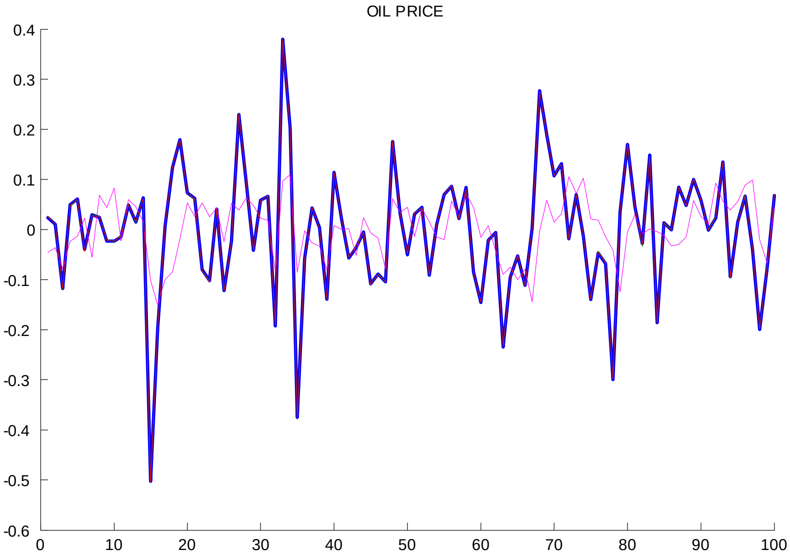
Using current shocks & 1 & 2 lags



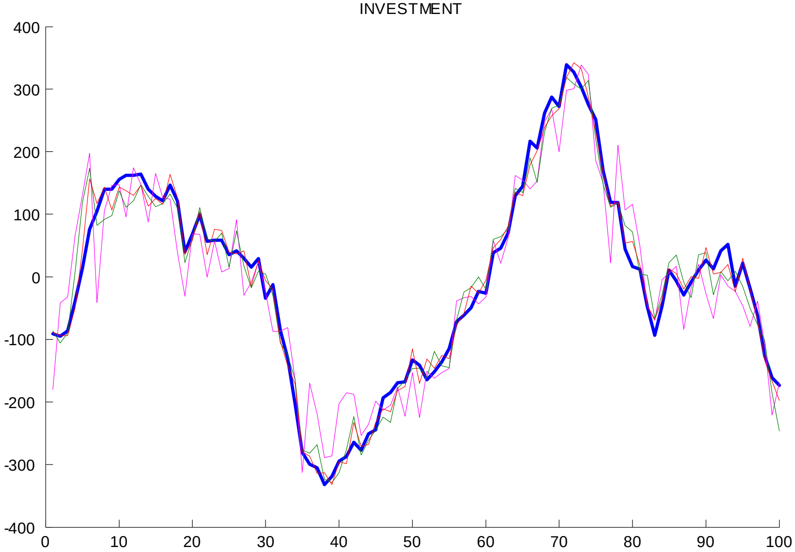
Using current shocks & 1 & 2 lags



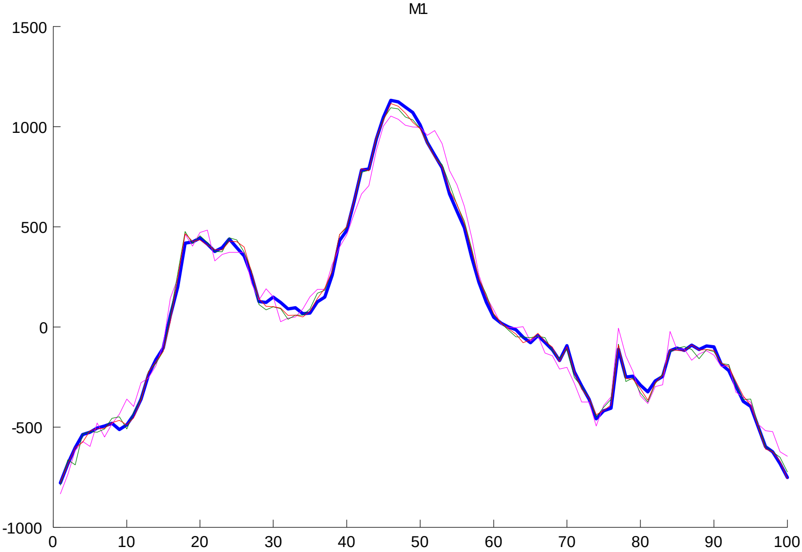
Using current shocks & 1 & 2 lags



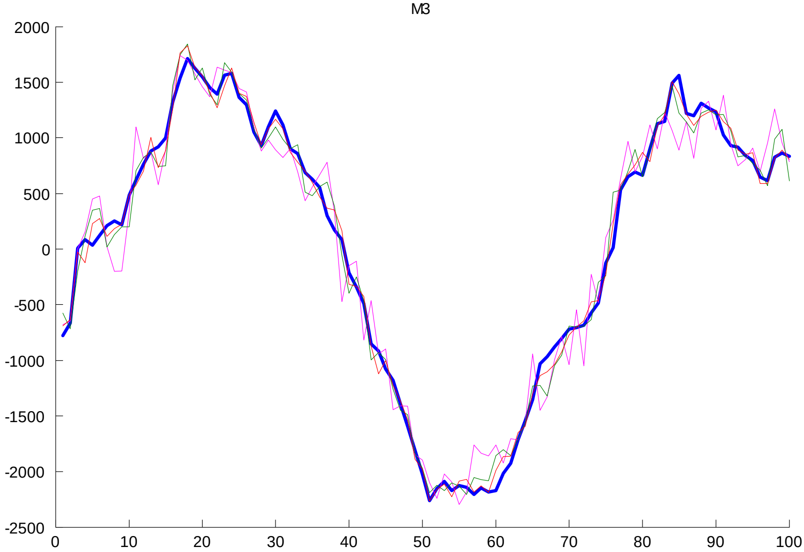
Using current shocks & 1 & 2 lags



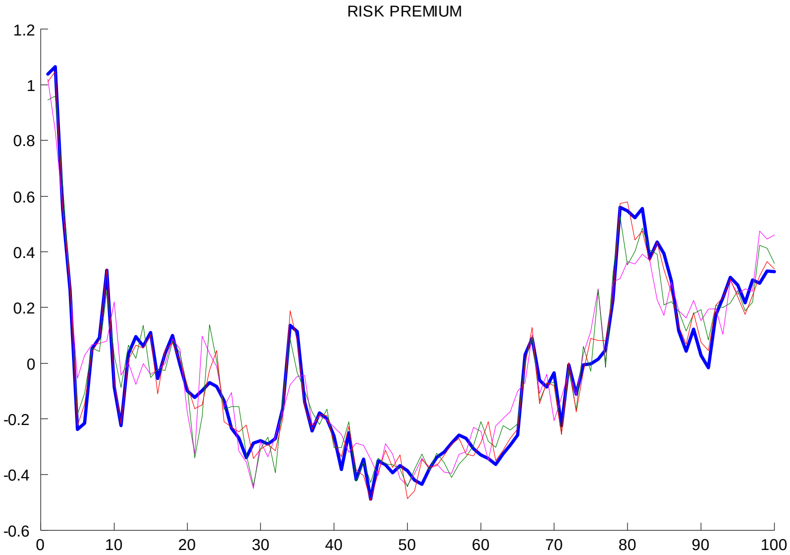
Using current shocks & 1 & 2 lags



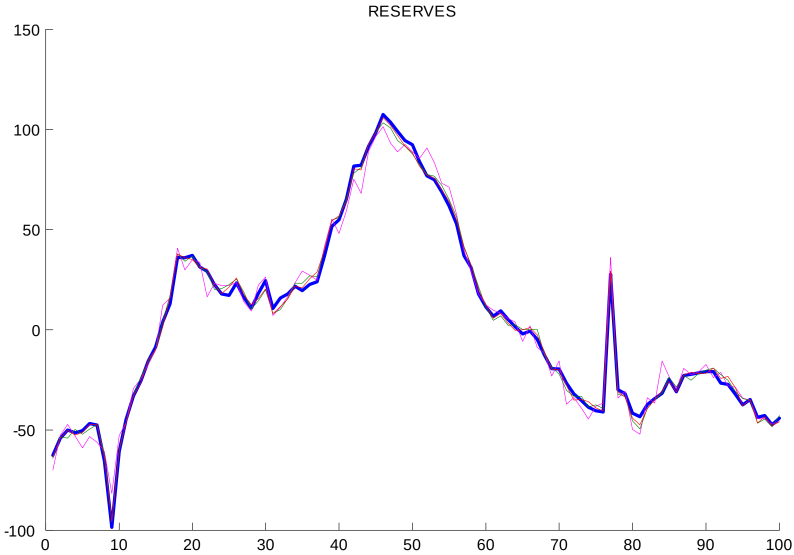
Using current shocks & 1 & 2 lags



Using current shocks & 1 & 2 lags



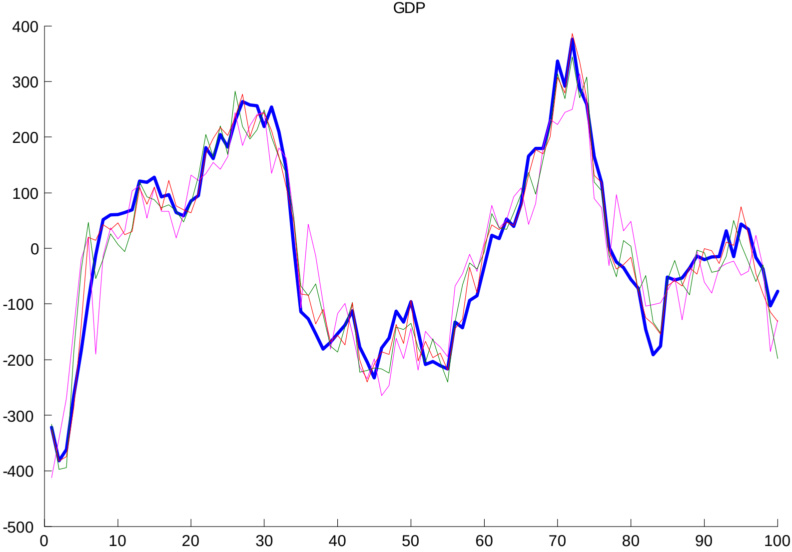
Using current shocks & 1 & 2 lags



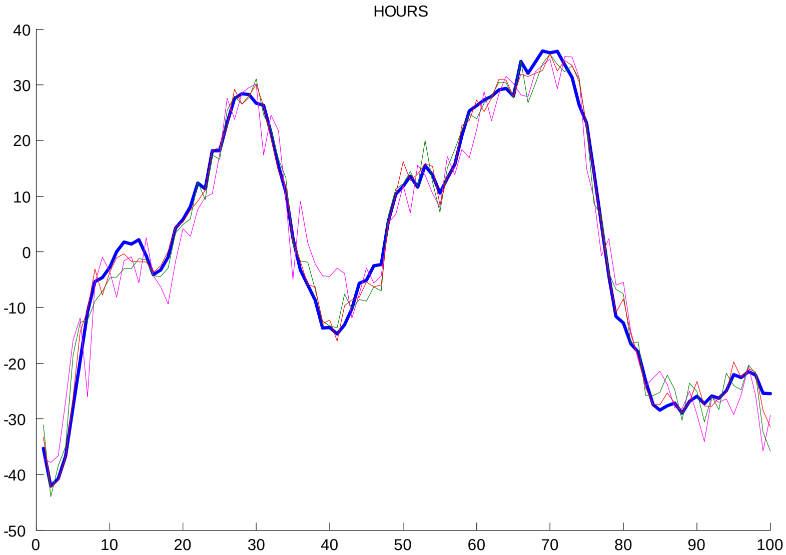
Using current shocks & 1 & 2 lags

Using current shocks & 1 & 2 lags

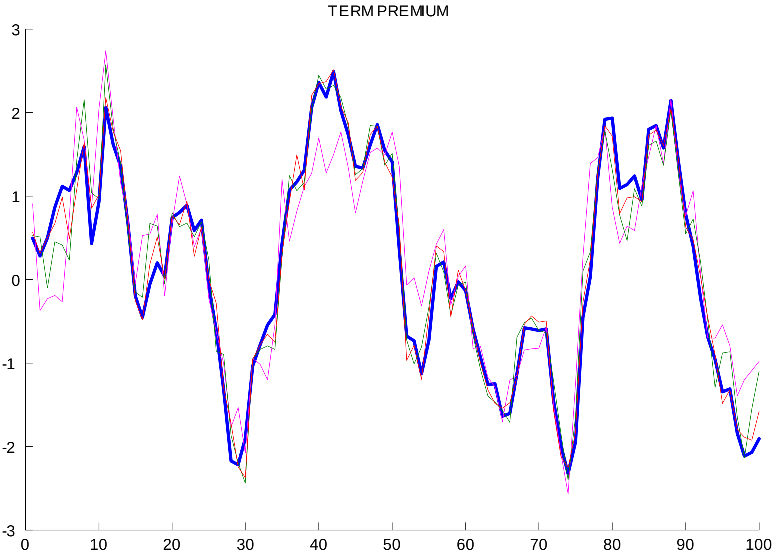
Using current shocks & 1 & 2 lags



Using current shocks & 1 & 2 lags



Using current shocks & 1 & 2 lags



R2

	Only current	+1 lags	+2 lags
Consumption	0.87	0.95	0.97
Credit	0.95	0.97	0.98
Inflation	**0.51**	0.81	0.90
Inflation_inv	**0.63**	1.0	1.0
Oil price	**0.21**	1.0	1.0
Investment	0.89	0.96	0.99
risk premium	0.86	0.93	0.98
hours in RBC	*0.825*	0.971	0.994

R2

	Only current	+1 lags	+2 lags
M1	0.98	1.00	1.00
M3	0.96	0.99	0.99
reserves	0.99	1.0	1.0
stock market	0.93	0.97	0.98
wage rate	0.94	0.98	0.99
GDP	0.83	0.92	0.97
hours	0.94	0.98	0.99
term premium	**0.80**	0.94	0.98
hours in RBC	*0.825*	0.971	0.994

Shocks versus the model

- Is this bad?
- Maybe not, but perceived wisdom—and the language in the paper—suggests that *propagation* is very important

Shocks versus the model

Explaining US vs EA

DSGE model:

$$z_{t+1}^{us} - z_{t+1}^{ea} = \tilde{a}_0 + \tilde{A}_2(shocks_t^{us} - shocks_t^{ea})$$

- !!! Use the same model for US and EA
- Only differences are the shocks?

R2

	Only current	+1 lags	+2 lags
Consumption	0.72	0.88	0.94
Credit	0.83	0.90	0.94
Inflation	**0.53**	0.66	0.76
Inflation_inv	**0.31**	1.0	1.0
Oil price	**0.23**	1.0	1.0
Investment	0.88	0.94	0.96
risk premium	0.86	0.92	0.95

R2

	Only current	+1 lags	+2 lags
M1	0.96	0.98	0.99
M3	0.75	0.87	0.92
stock market	0.75	0.87	0.92
wage rate	0.80	0.87	0.92
GDP	0.83	0.93	0.96
term premium	0.93	0.96	0.98

Very tricky issue

- Data have trends
- Methodology works with stationary data

Simple solution

- Put stochastic trend in model
- Use first differences of model variables

Disadvantages of the simple solution

- Loose level information
  - e.g., consumption $\approx 2/3$ output
  - $\Longrightarrow$ less parameters are identified
- Δ-filter emphasizes high frequency
  - measurement error shock could absorb this
- Not obvious how to put the right trend in model
- Not obvious data are consistent with balanced growth

Alternative I

Detrend data using

$$y_t = a_0 + a_1 t + a_2 t^2 + u_y$$

Advantage

- Each observable can have its own trend

Alternative II

Explicitly model trend

$$y_t^{\text{obs}} = y_t^{\text{trend}} + y_t^{\text{model}}$$

$$y_t^{\text{trend}} = \mu + y_{t-1}^{\text{trend}} + e_{y,t}$$

y_t^{model} is cyclical component determined by usual linearized equations

Advantage

- You can be very flexible in writing process for trend
- Different variables can have different trends

Alternative II

Same but written differently

$$
\begin{aligned}
y_t^{*\text{obs}} &= \Delta y_t^{\text{trend}} + \Delta y_t^{\text{model}} \\
\Delta y_t^{\text{trend}} &= \mu + e_{y,t}
\end{aligned}
$$

Alternative III ??

Detrend data using HP or Band-Pass filter

$$y_t^{\text{obs-filtered}} = B(L) y_t^{\text{obs}}$$

Problem

- $B(L)$ is a two-sided filter $\Longrightarrow$
- $y_t^{\text{obs-filtered}}$ has different properties than model data $\Longrightarrow$
- apply same filter do model data

$$y_t^{\text{model-filtered}} = y_t^{\text{model}}$$

Estimating misspecified models

- shocks versus observables
- would wedges work?

Alternatives to Bayesian estimation

- Maximum likelihood
- Calibration
- GMM
- SMM & indirect inference