Bayesian Estimation

Extensions

Tools for Macroeconomists: The essentials

Petr Sedláček

Extensions

Extensions

- Bayesian inference and model comparison
- Dealing with trends
- More on priors

Extensions

Bayesian Inference and Model Comparison

Bayesian vs. frequentist inference

- Bayesian inference cannot use frequentist principles
  - t-test, F-test, LR-test etc.
  - they have a frequentist justification of repeated sampling
- instead, there are two common Bayesian principles:
  - Highest Posterior Density (HPD) interval
  - Bayes factors (posterior odds)

Highest posterior density intervals

A $100(1-\alpha)\%$ posterior interval for Ψ is given by

$$P(\underline{b} < \Psi < \overline{b}) = \int_{\underline{b}}^{\overline{b}} P(\Psi|\mathcal{Y}^T)d\Psi = 1-\alpha$$

- there exists many such intervals
- the HPD interval is the smallest one of them

HPD "tests"

- the HPD test amounts to checking whether $\Psi_i \in HPD_{1-\alpha}$
- this is an informal way of comparing nested models
    - i.e. different parameter values
- Bayesians can also compare non-nested models
- more on this below

Bayes factors

$$B = \frac{P(\mathcal{Y}^T|\Psi_1)P(\Psi_1)}{P(\mathcal{Y}^T|\Psi_2)P(\Psi_2)}$$

- where Ψ_1 and Ψ_2 are two different sets of parameter values
- if $B > 1 \rightarrow \Psi_1$ is *a posteriori* more likely than Ψ_2
- important to use “proper” priors!

Model comparison

- posterior densities can be used to evaluate
  - conditional probabilities of particular parameter values
  - conditional probabilities of different model specifications
- use Bayes factors (posterior odds ratio) to compare models
  - advantage is that all models are treated symmetrically
  - there is no “null” model compared to an alternative

Model comparison

$$B_{A|B} = \frac{P_A(\mathcal{Y}^T|\Psi_A)P_A(\Psi_A)}{P_B(\mathcal{Y}^T|\Psi_B)P_B(\Psi_B)}$$

- it is also possible to assign priors on *models*
- the posterior odds ratio is then

$$PO_{A|B} = \frac{P(A|\mathcal{Y}^T)}{P(B|\mathcal{Y}^T)} = B_{A|B}\frac{P(A)}{P(B)}$$

How much information in Bayes factor?

Kass and Raftery (1995), if the value of $B_{A|B}$ is

- between 1 and 3 $\rightarrow$ barely worth mentioning
- between 3 and 20 $\rightarrow$ positive evidence
- between 20 and 150 $\rightarrow$ strong evidence
- over 150 $\rightarrow$ very strong evidence

Extensions

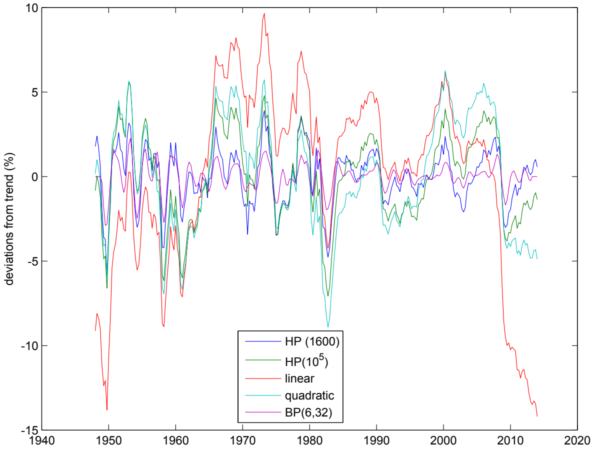
Trends

Trends

Problem:

- methodology works for stationary environments
- data has trends
- not clear which trend the model represents?

Trends

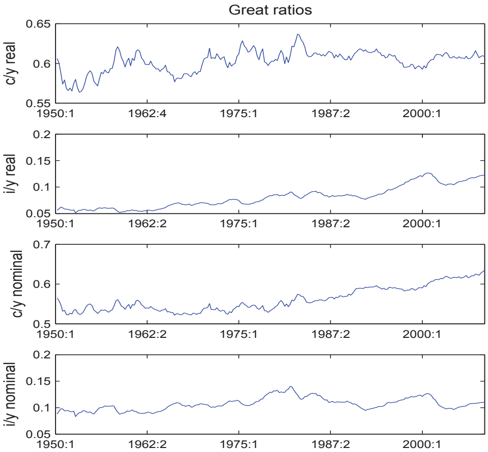


10
5
0
-5
-10
-15
deviations from trend (%)
1940
1950
1960
1970
1980
1990
2000
2010
2020
HP (1600)
HP(10^5)
linear
quadratic
BP(6,32)

Trends

- we could build in a trend within the model
- e.g. productivity is trending
- "stationarize" non-stationary variables within the model
- i.e. inspect variables relative to productivity
- however, not clear that data satisfies balanced growth

Trends



Great ratios
c/y real
0.65
0.6
0.55
1950:1
1962:4
1975:1
1987:2
2000:1
i/y real
0.2
0.15
0.1
0.05
1950:1
1962:2
1975:1
1987:2
2000:1
c/y nominal
0.7
0.6
0.5
1950:1
1962:2
1975:1
1987:2
2000:1
i/y nominal
0.2
0.15
0.1
0.05
1950:1
1962:2
1975:1
1987:2
2000:1

Trends

Solutions:

- use differenced data
  - highlights high-frequency movements (measurement error)
- detrend prior to estimation

Estimation on detrended data

- use e.g. quadratic trend:

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

  - each variable can have its own trend
- using HP or Band Pass filter:

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

  - $B(L)$ is a 2-sided filter!
  - $\rightarrow$ creates artificial serial correlation in the filtered data
  - $\rightarrow$ apply filter also to model data

Estimation on detrended data

- the above implies that the model is fitted to low(er) frequencies only
- Canova (2010) points out that the above can lead to:
    - underestimated volatility of shocks
    - persistence of shocks is overestimated
    - less perceived noise $\rightarrow$ policy rules imply higher predictability
    - substitution and income effects may be distorted due to above
- proposes to estimate flexible trend specifications within model

Extensions

More on priors

More on selecting priors

- so far selecting (independent) priors about deep parameters
- however, this may not be the best way to go
  - often we have priors about *observables*
  - good independent priors may still cause weird model properties
- solutions proposed in the literature:
  - Del Negro, Schorfheide (2008)
  - Andrle, Benes (2013)
  - Jarocinsky, Marcet (2013)

Del Negro, Schorfheide (2008)

- more guidance for eliciting priors
- issues with (independent) priors on deep parameters:
  - may lead to probability mass on unrealistic model properties
  - most exogenous shock processes are latent $\rightarrow$ what priors?
  - priors are often transfered to different models

Del Negro, Schorfheide (2008)

- they group parameters into three categories:
  - those determining the steady state
  - those determining exogenous shocks
  - those determining the endogenous propagation mechanism

Del Negro, Schorfheide (2008)

Parameters related to steady state relationships

- discount rate, depreciation, returns to scale etc.
- let $S_D(\Psi_{ss})$ be a vector of steady state relationships
  - Ψ_{ss}: set of parameters
- then $\widehat{S} = S_D(\Psi_{ss}) + \eta$ are their measurements
  - η: measurement error
- $\widehat{S}$ has a probabilistic interpretation and therefore
- using Bayes' rule, one can write $P(\Psi_{ss}|\widehat{S}) \propto \mathbb{L}(\widehat{S}|\Psi_{ss})P(\Psi_{ss})$
- allows for overidentification

Del Negro, Schorfheide (2008)

Exogenous processes

- volatility and persistence parameters
- use implied moments of endog. variables to "back out" priors
- the above is given values for Ψ_{ss} and Ψ_{endo}
- $\rightarrow$ valid for a particular model
- i.e. should not be directly transfered across models

Del Negro, Schorfheide (2008)

Endogenous propagation mechanisms

- price rigidity, labor supply elasticity etc.
- one could use similar principle as above
- authors suggest independent priors
    - researchers often have a relatively good idea
- joint prior induces non-linear relationships between parameters
- joint prior becomes

$$P(\Psi|\widehat{S}) \propto \mathbb{L}(\widehat{S}|\Psi_{ss})P(\Psi_{ss})P(\Psi_{endo})$$

- requires an additional step in MCMC algorithm

Andrle, Benes (2013)

- do not distinguish between groups of parameters
- their “system priors” are priors about concepts such as
  - impulse response functions
  - conditional correlations etc.

Andrle, Benes (2013)

- good independent priors may still cause weird model properties
- estimation can assign substantial mass on such regions
- call for careful prior-predictive analysis:
  - IRFs, second moments ...
  - compare with posterior results
  - is it the data or the model driving the results?

Andrle, Benes (2013)

Candidates for system priors:

- steady states
    - sensible values in levels or growth rates
- (un-)conditional moments
    - cross-correlations (conditional on shocks)
- impulse response properties
    - peak impacts, duration, horizon of MP effectiveness etc.

Andrle, Benes (2013)

Implementation:

- use Bayes' rule again
- specify model properties you care about $Z = h(\Psi)$
- these can be characterized by a probabilistic model $Z \sim D(Z^s)$
  - $D(Z^s)$ is a distribution function
  - Z^s are parameters of that function (hyper-parameters)
- its likelihood function (the system prior): $P(Z^s|\Psi, h)$
- composite joint prior: $P(\Psi|Z^s, h) \propto P(Z^s|\Psi, h)P(\Psi)$

Andrle, Benes (2013)

The posterior becomes

$$P(\Psi|\mathbb{Y}^T, Z^S) \propto \mathbb{L}(\mathbb{Y}^T|\Psi)P(Z^S|\Psi, h)P(\Psi)$$

- evaluation is in principle the same as before
- use of MCMC methods
- additional step in evaluating the system prior
- slows things down - have to run MCMC on prior

Jarocinsky, Marcet (2013)

- similar ideas as above, but in the context of Bayesian VARs
  - their point is that widely used priors about parameters
  - can lead to behavior of observables that is counterfactual
  - $\rightarrow$ always a good to do prior-predictive analysis of you model!