Econometric Policy Analysis under Uncertainty

Sargan Lectures of the Econometric Society - Sydney 2023

Raffaella Giacomini (University College London and Chicago Fed)[1]

[1] The lectures do not reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System

Introduction

What uncertainty?

- Economists have grappled with uncertainty since the dawn of the discipline (e.g., Knight, 1921; Keynes, 1921)
- Different types of uncertainty:
  1. Aleatory uncertainty = things we cannot know (e.g., future shocks)
     - ⋆ We are relatively good at this, think of reporting forecasts as probability distributions (e.g., Bank of England forecasts of inflation)
  2. Epistemic or scientific uncertainty = things we could in principle know (e.g., effect of past policies) but have limited knowledge about due to uncertainties of the scientific process
     - ⋆ Scholars in different fields have argued that we are pretty bad at reporting this (Manski 2019's "lure of incredible certitude")
     - ⋆ Side note: interesting research question is how experts (do and should) communicate scientific uncertainty (Van der Bles et al., 2019).

Dangers of overstating certainty

- Overstating certainty dangerous when science informs policy
  - Famous example: US and UK declaring war in Iraq due to reports of non-existent "weapons of mass destruction" by the intelligence community
  - Later investigations found that reports "did not accurately explain to policymakers the uncertainties behind the judgments"
- Economists in the spotlight due to direct influence on economic policy decisions
- Here we focus on the more benign case of econometric analysis that affects policy decisions via estimation of causal effects:
  1. Dynamic causal effects of shocks (impulse-response analysis)
  2. Average treatment effects of a policy intervention

Uncertainties in causal effects estimation

- We will consider different kinds of scientific uncertainties
  - Impulse-response analysis
    - ⋆ Uncertainty about identifying assumptions (Lecture 1)
    - ⋆ Uncertainty about the measurement of shocks (Lecture 2)
  - Average treatment effects (Lecture 3)
    - ⋆ Uncertainty about the validity of the control group
    - ⋆ Uncertainty about model specification

Uncertainty and robustness

- Here we propose to deal with uncertainty in econometric methods using the notion of robustness
- Different ways to see robustness
  - Drop assumptions you are uncertain about and see what you can still say $\rightarrow$ look for robust econometric methods = valid under minimal assumptions (Lectures 1-3)
  - Give a new estimator that incorporates uncertainty about assumption $\rightarrow$ robust decision under uncertainty (Lecture 1)
  - Understand impact of assumptions you are uncertain about $\rightarrow$ formal sensitivity analysis (Lecture 1)

Lecture 1. Uncertainty about identification of shocks

Based on:

- Giacomini, R. and T. Kitagawa (2021), "Robust Bayesian analysis for set-identified models"
- Giacomini, R., T. Kitagawa and H. Ulhig (2019), "Estimation under ambiguity"

A shock is in the eyes of the beholder

- We focus on impulse-response analysis in Structural Vector Autoregressive models (SVARs)
- In order to infer causation from observed correlations we need to impose identifying assumptions
- Typically, we impose enough assumptions to be able to point-identify the structural parameters of interest (the impulse responses)
- The profession does not agree on these assumptions (some are particularly controversial) $\rightarrow$ uncertainty about identifying assumptions
- Implies that we cannot compare studies using different identifying assumptions, because their definition of shock is different

What to do about uncertain identification?

- In a nutshell, lecture 1 will present two approaches to impulse-response analysis under uncertain identification

1. **Robust method:** Drop controversial identifying assumptions but then report impulse responses that are sets, not points
   - Pros: you can compare different models/studies, results truly reflect what's in the data+ plausible assumptions, not arbitrary assumptions
   - Cons: sets are typically wide and suggest no effect ("too much" uncertainty)

What to do about uncertain identification?

2. **Method that incorporates uncertainty:** Instead of reporting a set, report a 'robust decision under uncertainty', i.e., a point that minimizes the maximum loss over the set
   - Pros: easier to communicate point estimates rather than sets
   - Cons: requires some arbitrary choices

1. Robust method (Giacomini and Kitagawa, 21)

Motivation

- Empirical literature deals with uncertain identification by imposing "weak" sign restrictions in SVARs or dropping some equality restrictions $\rightarrow$ set-identified impulse response
- The literature does not however report an estimate of the set, but selects a point within the set (Uhlig, 05)
- How is the point chosen? Literature uses Bayesian inference and reports the posterior mean that corresponds to a particular prior
- Prior choice is arbitrary, see critique by Baumeister and Hamilton (15), who propose a way to choose prior more carefully

Why is this a problem?

- We normally don't worry too much about prior sensitivity, because under point identification the effect of the prior disappears asymptotically (know as the 'Bernstein von Mises property')
- This is not true anymore under set-identification (Poirier, 98)
- $\rightarrow$ Existing approach is not robust because it is driven by prior choice
- We want to propose a truly robust method, that is, robust not only to uncertain identification but also to prior choice

What to do?

- Trying a few different priors is NOT the solution (although often seen in applied literature)
- Solution: formal Bayesian approach that is robust to the choice of prior $\rightarrow$ formal sensitivity analysis
  - Giacomini and Kitagawa (21): do not choose prior and estimate the set
  - Giacomini, Kitagawa and Uhlig (19): choose prior but perturb it using a robust control approach
  - Giacomini, Kitagawa and Volpicella (22): choose prior but perturb it using a model averaging approach (won't talk about this)
- Key idea: shift from standard Bayesian (= single prior) to robust Bayesian (= multiple priors)

Example

- Static SVAR model for wage and employment growth, $x_t = (\Delta w_t, \Delta n_t)$

$$Ax_t = \epsilon_t,$$

$$A = \begin{pmatrix} -\beta_d & 1 \\ -\beta_s & 1 \end{pmatrix}, \ \epsilon_t \sim \mathcal{N}(0, I)$$

- Structural parameter: $\theta = (\beta_d, \beta_s)$

- (Scalar) parameter of interest: α: either one of the elasticities in θ or impulse response = each element of A^{-1}

- Reduced-form parameter: $\phi = vec(\Sigma)$, $\Sigma = Var(x_t)$

Example

- Sign restrictions: downward sloping demand $\beta_d \leq 0$, upward sloping supply $\beta_s \geq 0$
- So parameter of interest α is set-identified
- Existing Bayesian approaches do not report a set for α, but select a point within the set

Single-prior approach of Uhlig (05)

- Choose prior for
  - reduced form parameter ϕ
  - orthonormal matrix Q spanning set of observationally equivalent θ's given ϕ:
    - Since $\Sigma = \Sigma_{tr} Q Q' \Sigma_{tr}' = A^{-1} A^{-1\prime}$ we have $A^{-1} = \Sigma_{tr} Q$
- Draw ϕ from its posterior (because this prior is updated) and Q from its prior (because this prior is not updated) and keep draws that satisfy the identifying restrictions
- Report posterior mean and credible region for α
- Critique (Baumeister and Hamilton, 05): uninformative prior for Q is spuriously informative for α

Single-prior approach of Baumeister and Hamilton (15)

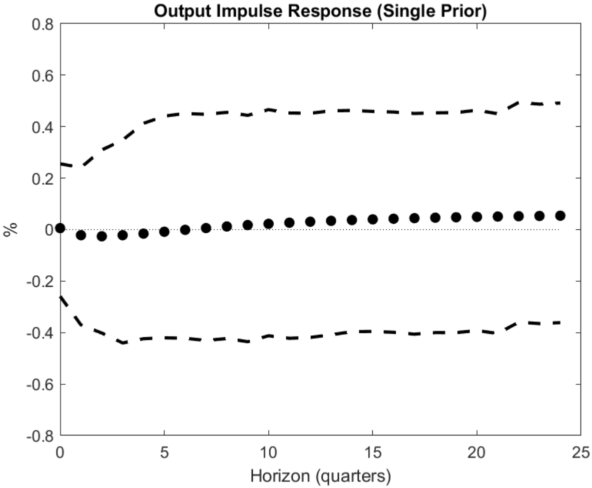
- Specify a prior directly for θ, e.g., based on micro-estimates:
- $\pi_{(\beta_d,\beta_s)}$: independent truncated Student's t such that $\pi_{\beta_s}([0.1, 2.2]) = 0.9$ and $\pi_{\beta_d}([-2.2, -0.1]) = 0.9$
- Critique: Shape of prior arbitrary, inference remains driven by prior choice

Multiple-prior approach of Giacomini Kitagawa (21)

- Idea:
  - ϕ and Q are different: data informative about ϕ but not Q
  - $\rightarrow$ choose single prior for ϕ but arbitrary multiple priors for Q
  - In practice, do the following: for each draw from posterior for ϕ, draw many orthonormal Q's, discard draws that do not satisfy restrictions and then compute upper and lower bound for α
  - Gives draws of upper and lower bounds, from which can construct "set of posterior means" (estimator of identified set) and "robust credible region" (uncertainty about identified set estimator)
- Eliminates prior sensitivity and restores asymptotic equivalence between Bayesian and frequentist inference under set-identification
- Critique: Identified set estimator may be too wide and set of priors may include unrealistic ones

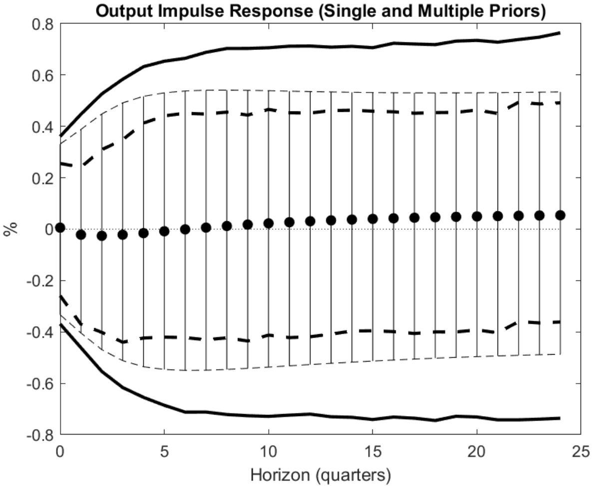
Single- vs. multiple-prior for impulse response analysis

Uhlig (05)



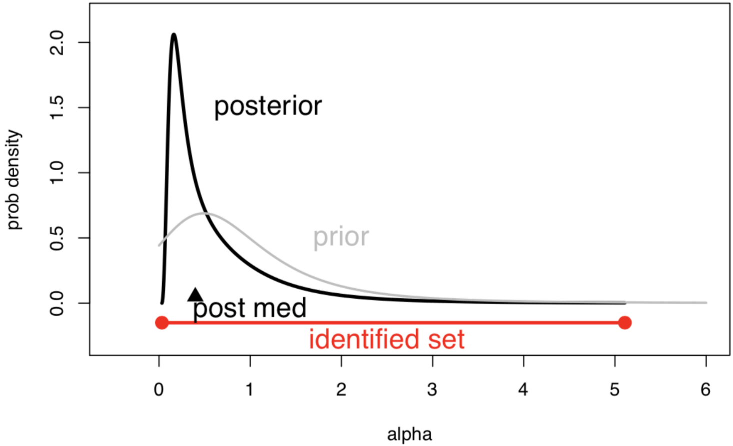
Single- vs. multiple-prior for impulse response analysis

Uhlig (05) vs. Giacomini Kitagawa (21)



Single- vs. multiple-prior for elasticity estimation

Baumeister and Hamilton (15) vs. Giacomini Kitagawa (21)



What have we learned?

- In impulse-response analysis we are uncertain about identifying restrictions
- Can drop the problematic restrictions or use weaker sign restrictions but then need to report a set, not an arbitrary point in the set
- In other words, when you impose sign restrictions but report a point, your inference is driven by the choice of an unrevisable prior
- Instead, consider a robust Bayesian approach (=multiple prior) that eliminates effect of prior choice and gives set estimates
- Have to live with the fact that these sets are large. Next approach tries to do something about this...

2. Method that incorporates uncertainty (Giacomini, Kitagawa and Uhlig, 19)

The idea

- Same setup with weak assumptions that give set-identification and thus sensitivity to prior choice if you want to be Bayesian
- Want to report a point instead of a set, but how to choose it non-arbitrarily? $\rightarrow$ Find a robust estimator
- Choose a benchmark prior π^* and perturb it by considering the set of priors in its Kullback-Leibler (KL) neighborhood
- Do minmax estimation over resulting set of posteriors
- Similar to the robust control methods of Hansen & Sargent (01) for decisions under misspecification concerns. Here:
  - decision = estimator
  - misspecification concern is about the prior

Which benchmark prior?

- Starting point is to express a joint prior for (θ, ϕ) as $\pi^*_{\theta|\phi}\pi^*_{\phi}$, where
  - π^*_{ϕ} revisable $\rightarrow$ keep fixed
  - $\pi^*_{\theta|\phi}$ unrevisable $\rightarrow$ perturb

KL neighborhood of the benchmark prior $\pi^*_{\theta|\phi}$

- For radius $\lambda > 0$ consider all priors in the KL-neighborhood of $\pi^*_{\theta|\phi}$:

$$KL = \left\{ \pi_{\theta|\phi} : \int \ln \left(\frac{d\pi_{\theta|\phi}}{d\pi^*_{\theta|\phi}} \right) d\pi_{\theta|\phi} \leq \lambda \right\}$$

- Need to choose radius λ (= degree of uncertainty), more on this later...

The set of priors and posteriors for α

- Henceforth, priors in red and posteriors in blue
- The KL neighborhood implies a set of priors for scalar parameter of interest $\alpha = \alpha(\theta, \phi)$:

$$\Pi_\alpha^\lambda = \left\{ \pi_\alpha = \int \pi_{\theta|\phi}(\alpha(\theta,\phi)) \, d\pi_\phi^* : \pi_{\theta|\phi} \in KL \right\}$$

- ...and a set of posteriors, by updating only the reduced-form prior:

$$\Pi_{\alpha|X}^\lambda = \left\{ \pi_{\alpha|X} = \int \pi_{\theta|\phi}(\alpha(\theta,\phi)) \, d\pi_{\phi|X}^* : \pi_{\theta|\phi} \in KL \right\}$$

- We want a point estimator $\widehat{\alpha}$ of α

Estimation as a minmax decision under uncertainty

- $\widehat{\alpha}$: decision and $L(\widehat{\alpha}, \alpha)$: loss function. E.g., $L(\widehat{\alpha}, \alpha) = (\widehat{\alpha} - \alpha)^2$
- The estimator minimizes the worst-case posterior expected loss:

$$\min_{\widehat{\alpha}} \int \left[\max_{\pi_{\theta|\phi} \in KL} \int L(\widehat{\alpha}, \alpha) \; d\pi_{\theta|\phi} \right] d\pi_{\phi|X}$$

- Can rewrite in terms of the worst-case prior π^0:

$$\min_{\widehat{\alpha}} \int \int L(\widehat{\alpha}, \alpha) \; d\pi^0_{\alpha|\phi} \; d\pi^*_{\phi|X},$$

$$d\pi^0_{\alpha|\phi} \propto \exp L(\widehat{\alpha}, \alpha)/\kappa^* d\pi^*_{\alpha|\phi} \,,$$

with κ^* solution to

$$\min_{\kappa \geq 0} \left\{ \kappa \ln \int \exp \left\{ \frac{L(\widehat{\alpha}, \alpha)}{\kappa} \right\} d\pi^*_{\alpha|\phi} + \kappa\lambda \right\}$$

Obtaining the minimax estimator in practice

- Can compute minimax estimator as long as we can:
  - Draw from the reduced-form parameter posterior $\pi^*_{\phi|X}$
  - Draw from or evaluate the benchmark prior for α, $\pi^*_{\alpha|\phi}$

The Baumeister and Hamilton (15) example

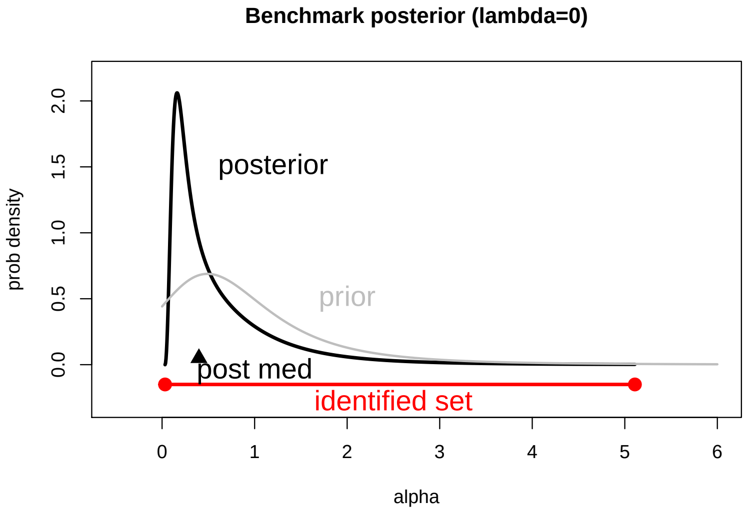
- SVAR

$$A_0 \begin{pmatrix} \Delta n_t \\ \Delta w_t \end{pmatrix} = c + \sum_{k=1}^{8} A_k \begin{pmatrix} \Delta n_{t-k} \\ \Delta w_{t-k} \end{pmatrix} + \begin{pmatrix} \epsilon_t^d \\ \epsilon_t^s \end{pmatrix},$$

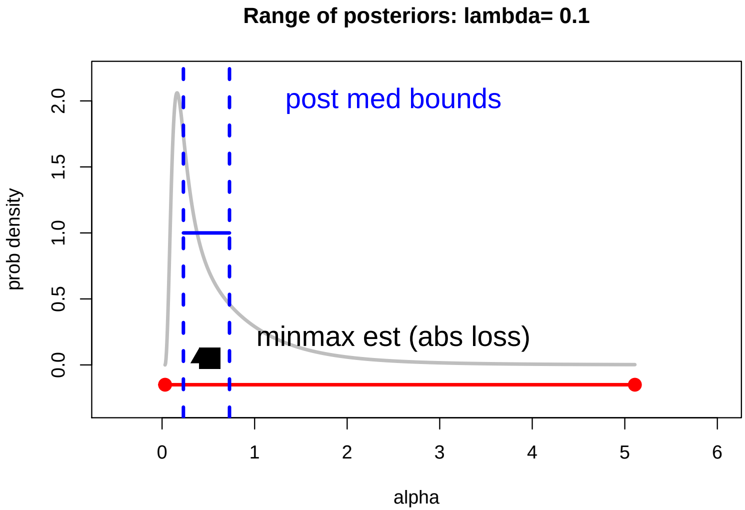
$$A_0 = \begin{pmatrix} -\beta_d & 1 \\ -\beta_s & 1 \end{pmatrix}, \ (\epsilon_t^d, \epsilon_t^s) \sim \mathcal{N}(0, diag(d_1, d_2)).$$

- Say parameter of interest is $\alpha = \beta_s$

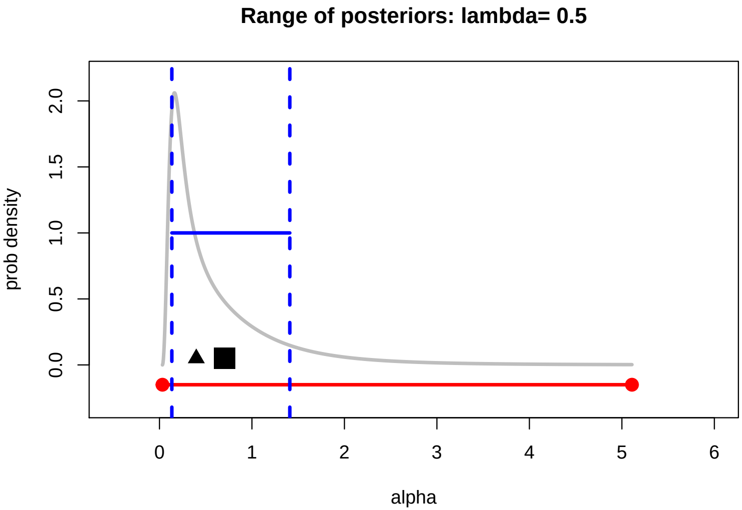
Consider the Baumeister and Hamilton (15) prior as the benchmark



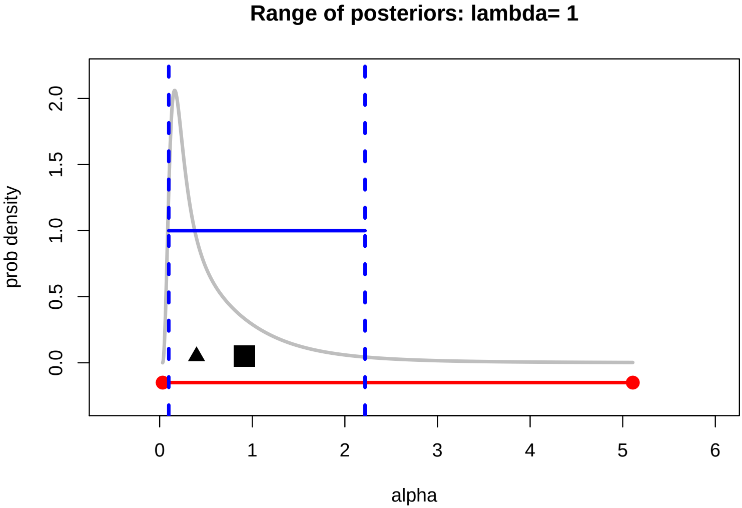
Minmax estimation for absolute loss, given λ



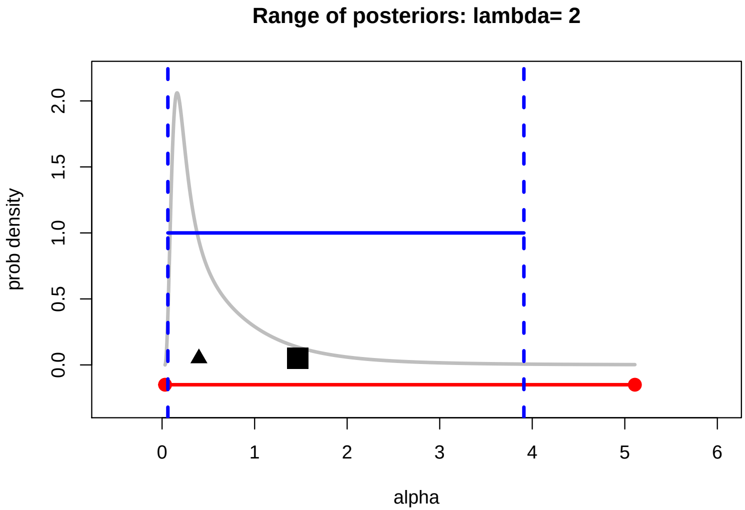
Minmax estimation for absolute loss, given λ



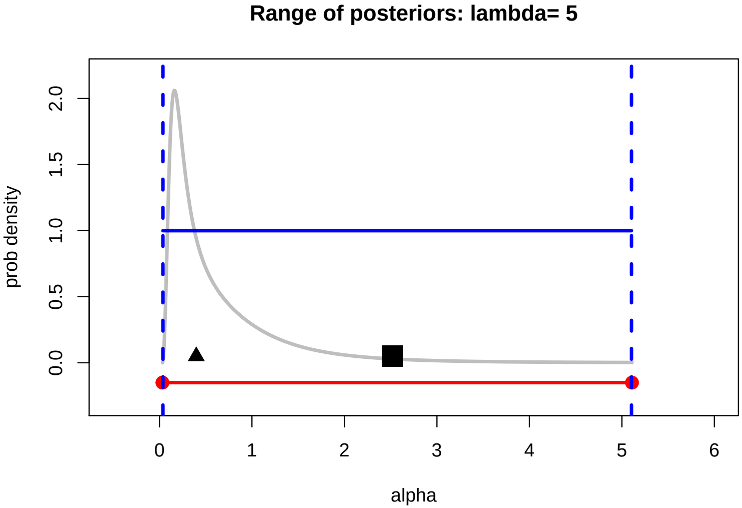
Minmax estimation for absolute loss, given λ



Minmax estimation for absolute loss, given λ



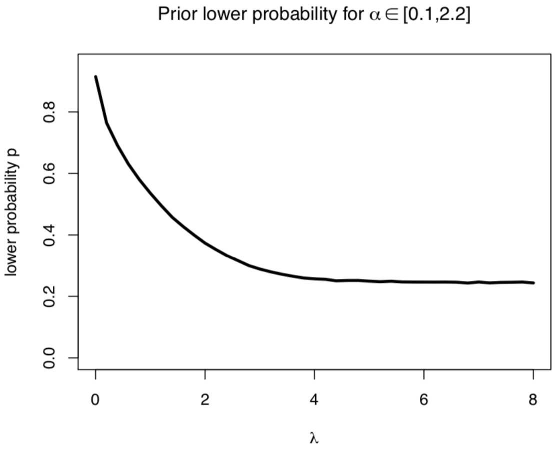
Minmax estimation for absolute loss, given λ



How to choose the degree of uncertainty parameter λ?

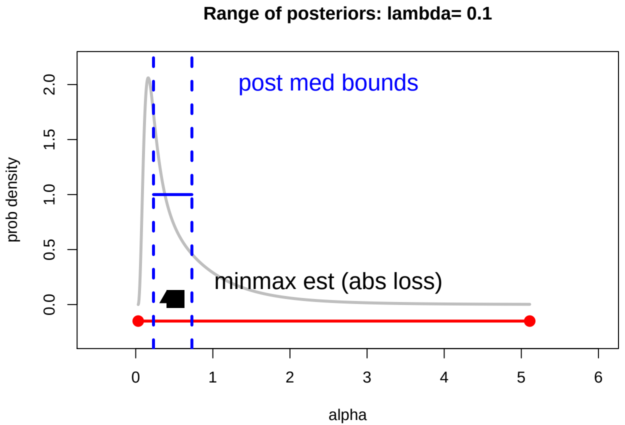
- Known challenge in robust control: λ has no interpretable scale
- Idea: map candidate values for λ into a set of priors for a parameter for which we have (partial) prior knowledge
  - Can do this because we show invariance to reparameterization
- Then choose the λ that best fits our prior knowledge
- Example: say we have a prior on the probability that α lies in a given range (e.g., $\pi_\alpha([0.1, 2.2]) = .9$ in BH)

Probability of $\alpha \in [0.1,2.2]$ for different λ's



So if our prior is that this is 0.9, we want to choose small λ

So this may be considered the preferred estimator



So our robust estimator (square) is a little larger than but not that different from Baumeister and Hamilton (15)'s (triangle)

Conclusion

- We have considered impulse-response analysis using SVARs subject to identifying assumptions
- Here robustness to uncertain identification = considering weaker assumptions, which creates set-identification
- Communicating this uncertainty means reporting sets, not points
- The applied literature (Bayesian) reports points, which introduces sensitivity to prior choice that doesn't go away
- Showed two ways to robustify against prior choice
  1. Consider multiple priors for the unrevisable component of the prior and report estimator of identified set
  2. Perturb the unrevisable component of the prior and report the minmax estimator (point)

References

- Baumeister, C. and J. Hamilton (2015): Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information, Econometrica
- Giacomini, R., T. Kitagawa, and H. Uhlig (2019): Estimation Under Ambiguity, cemmap Working Paper CWP24/19.
- Giacomini, R. and T. Kitagawa (2021), Robust Bayesian inference for set-identified models, Econometrica
- Giacomini, R., T. Kitagawa and A. Volpicella (2022), Uncertain identification, Quantitative Economics
- Hansen, L. P. and T. J. Sargent (2001): Robust Control and Model Uncertainty, American Economic Review, AEA Papers and Proceedings,
- Knight FH. (1921) Risk, uncertainty, and profit. Chicago, IL: University of Chicago Press

References

- Keynes JM. (1921) A treatise on probability. London, UK: Macmillan
- Manski, C. F. (2019). The lure of incredible certitude (pp. 1â30). Economics and Philosophy
- Poirier, D. (1998): Revising Beliefs in Nonidentified Models, Econometric Theory
- Uhlig, H. (2005): What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure, Journal of Monetary Economics
- Van der Bles, A. M., van der Linden, S., AlexandraFreeman, L. J., Mitchell, J., Galvao, A. B., Zaval, L., Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. Royal Society Open Science,