Estimation under Ambiguity

R. Giacomini (UCL), T. Kitagawa (UCL), H. Uhlig (Chicago)
Questions:

- How to perform posterior analysis (inference/decision) when a probabilistic prior is not available or not credible?
- How to assess robustness of the posterior results to a choice of prior?
Robust Bayes viewpoint

- **I.J. Good** (p5, 1965, “The Estimation of Probabilities”): “My own view, following Keynes and Koopman, is that judgements of probability inequalities are possible but not judgements of exact probabilities; therefore a Bayesian should have upper and lower betting probabilities.”
Robust Bayes viewpoint

- **I.J. Good** (p5, 1965, “The Estimation of Probabilities”): “My own view, following Keynes and Koopman, is that judgements of probability inequalities are possible but not judgements of exact probabilities; therefore a Bayesian should have upper and lower betting probabilities.”

- Prior input: **a set of priors (ambiguous belief)**

- An updating rule is applied to the ambiguous belief

- Posterior output: **the set of posteriors** and **decision under ambiguity**
Example: SVAR

- Baumeister & Hamilton (15 Ecta, BH hereafter) two-variable SVAR(8), \((\Delta n_t, \Delta w_t)\): employment and wages growth

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

\[
A_0 = \begin{pmatrix} -\beta_d & 1 \\ -\beta_s & 1 \end{pmatrix}, \quad (\epsilon_t^d, \epsilon_t^s) \sim \mathcal{N}(0, \text{diag}(d_1, d_2)).
\]
Example: SVAR

- Baumeister & Hamilton (15 Ecta, BH hereafter) two-variable SVAR(8), \((\Delta n_t, \Delta w_t)\): employment and wages growth

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

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

- Parameter of interest: supply/demand elasticities and structural impulse responses.
- Bayesian practice in sign-restricted SVARs
  - Agnostic prior: Canova & De Nicolo (02 JME), Faust (98 CRSPP), Uhlig (05 JME)
  - Carefully elicited prior: BH
Prior elicitation procedure recommended by BH:

- Joint prior of \((\beta_d, \beta_s, d_1, d_2, A)\):

\[
\pi(\beta_d, \beta_s) \cdot \pi(d_1, d_2 | (\beta_d, \beta_s)) \cdot \pi A | (\beta_d, \beta_s, d_1, d_2)
\]

- \((\beta_d, \beta_s)\): Independent truncated t-dist. s.t. \(\pi_{\beta_s}([0.1, 2.2]) = 0.9\) and \(\pi_{\beta_d}([-2.2, -0.1]) = 0.9\).

- \((d_1, d_2) | (\alpha, \beta)\): Independent inverse gamma s.t. scales determined by \(A_0 \hat{\Omega} A'_0\).

- \(A | (\beta_d, \beta_s, d_1, d_2)\): Independent Gaussian priors shrinking the reduced-form coefficients \(A_0^{-1} A\) to random walks.
Prior elicitation procedure recommended by BH:

- Joint prior of \((\beta_d, \beta_s, d_1, d_2, A)\):

\[
\pi(\beta_d, \beta_s) \cdot \pi(d_1, d_2 | (\beta_d, \beta_s)) \cdot \pi(A | (\beta_d, \beta_s, d_1, d_2))
\]

- \((\beta_d, \beta_s)\): Independent truncated t-dist. s.t. \(\pi_{\beta_s}([0.1, 2.2]) = 0.9\) and \(\pi_{\beta_d}([-2.2, -0.1]) = 0.9\).
- \((d_1, d_2) | (\alpha, \beta)\): Independent inverse gamma s.t. scales determined by \(A_0 \hat{\Omega} A_0'\).
- \(A | (\beta_d, \beta_s, d_1, d_2)\): Independent Gaussian priors shrinking the reduced-form coefficients \(A_0^{-1}A\) to random walks.

**Issues**

- Available prior knowledge fails to pin down the exact shape.
- Why conjugate priors?
- Why priors independent?
Limited credibility on prior causes concern of robustness, which is even severer due to set-identification.
- Limited credibility on prior causes concern of robustness, which is even severer due to set-identification

**Existing robust Bayes procedure:**

- Giacomini & Kitagawa (18): introduce arbitrary priors to unrevisable part of the prior, and draw posterior inference with the set of posteriors = posterior inference for identified set as in Kline & Tamer (16, QE), Chen, Christensen, & Tamer (19, Ecta), among others.

- Drawback: The set of priors is huge and may well contain unrealistic ones, e.g., an extreme scenario receives (subjective) probability one.
This paper

Our proposal: Robust Bayesian methods with a refined set of priors
This paper

**Our proposal**: Robust Bayesian methods with a refined set of priors

- Introduce a set of priors as a Kullback-Leibler (KL) neighborhood of a benchmark prior.
- Apply the Bayes rule to each prior and conduct the posterior analysis (inference and decision) with the resulting set of posteriors.
Our proposal: Robust Bayesian methods with a refined set of priors

- Introduce a set of priors as a Kullback-Leibler (KL) neighborhood of a benchmark prior.
- Apply the Bayes rule to each prior and conduct the posterior analysis (inference and decision) with the resulting set of posteriors
- Similar to the robust control methods of Peterson, James & Dupuis (00 IEEE) and Hansen & Sargent (01 AER), here applied to statistical decision problems with concerns about prior misspecification.
Consider two-variable SVAR(0) model:

\[ A_0 \begin{pmatrix} \Delta n_t \\ \Delta w_t \end{pmatrix} = \begin{pmatrix} \epsilon^d_t \\ \epsilon^s_t \end{pmatrix}, \quad A_0 = \begin{pmatrix} -\beta^d & 1 \\ -\beta^s & 1 \end{pmatrix}. \]

- structural shocks \((\epsilon^d_t, \epsilon^s_t) \sim \mathcal{N}(0, \text{diag}(d_1, d_2))\).
Consider two-variable SVAR(0) model:

\[ A_0 \begin{pmatrix} \Delta n_t \\ \Delta w_t \end{pmatrix} = \begin{pmatrix} \epsilon^d_t \\ \epsilon^s_t \end{pmatrix}, \quad A_0 = \begin{pmatrix} -\beta^d \\ -\beta^s \\ 1 \end{pmatrix}. \]

- Structural shocks \((\epsilon^d_t, \epsilon^s_t) \sim \mathcal{N}(0, \text{diag}(d_1, d_2))\).
- Structural parameters: \(\theta = (\beta^d, \beta^s, d_1, d_2) \in \Theta\)
- Parameter of interest: denoted by \(\alpha = \alpha(\theta) \in \mathbb{R}\), e.g., elasticity, impulse response of \(\Delta n_t\) to unit \(\epsilon^d_t\), etc.
- Reduced-form parameters: \(\Sigma_{tr}\) Cholesky decomposition of \(\Sigma\) the Var-cov matrix of \((\Delta n_t, \Delta w_t)\), \(\phi = (\sigma_{11}, \sigma_{21}, \sigma_{22}) \in \Phi\).
Set-identifying assumption: downward sloping demand $\beta_d \leq 0$ and upward sloping supply $\beta_s \geq 0$

The identified set for the labor supply elasticity (Leamer (81 REStat)), $\alpha = \beta_s$,

$$IS_{\alpha}(\phi) \equiv \begin{cases} 
[\sigma_{21}/\sigma_{11}, \sigma_{22}/\sigma_{21}], & \text{for } \sigma_{21} > 0, \\
[0, \infty), & \text{for } \sigma_{21} \leq 0.
\end{cases}$$
Set-identifying assumption: downward sloping demand $\beta_d \leq 0$ and upward sloping supply $\beta_s \geq 0$

The identified set for the labor supply elasticity (Leamer (81 RESStat)), $\alpha = \beta_s$,

$$IS_\alpha(\phi) \equiv \begin{cases} \left[\frac{\sigma_{21}}{\sigma_{11}}, \frac{\sigma_{22}}{\sigma_{21}}\right], & \text{for } \sigma_{21} > 0, \\ [0, \infty), & \text{for } \sigma_{21} \leq 0. \end{cases}$$

Starting from a prior $\pi_\theta$ for $\theta$, the posterior for $\alpha(\theta)$ is

$$\pi_{\alpha|X}(\cdot) = \int_{\Phi} \pi_{\theta|\phi}(\alpha(\theta) \in \cdot) d\pi_{\phi|X}$$
Before observing data

**Innovation**: Introduce the following set of priors: for $\lambda > 0$,

$$
\Pi^\lambda_\theta = \left\{ \pi_\theta = \int_\Phi \pi_{\theta|\phi} d\pi_\phi : \pi_{\theta|\phi} \in \Pi^\lambda(\pi_{\theta|\phi}^*) \forall \phi \right\},
$$

where $\pi_{\theta|\phi}^*$ is a benchmark conditional prior s.t. $\pi_{\theta|\phi}^*(IS_\theta(\phi)) = 1$, $\forall \phi$.

$$
\Pi^\lambda(\pi_{\theta|\phi}^*) = \left\{ \pi_{\theta|\phi} : \mathcal{R}(\pi_{\theta|\phi} \| \pi_{\theta|\phi}^*) \leq \lambda \right\},
$$

$$
\mathcal{R}(\pi_{\theta|\phi} \| \pi_{\theta|\phi}^*) = \int_\Theta \ln \left( \frac{d\pi_{\theta|\phi}}{d\pi_{\theta|\phi}^*} \right) d\pi_{\theta|\phi}: \text{KL-distance relative to the benchmark prior.}
$$
Before observing data

**Innovation**: Introduce the following set of priors: for $\lambda > 0$,

$$\Pi^\lambda_\theta = \left\{ \pi_\theta = \int_\Phi \pi_{\theta|\phi} d\pi_\phi : \pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_\theta|\phi) \forall \phi \right\},$$

where $\pi^*_\theta|\phi$ is a benchmark conditional prior s.t. $\pi^*_\theta|\phi(IS_\theta(\phi)) = 1, \forall \phi$.

$$\Pi^\lambda(\pi^*_\theta|\phi) = \left\{ \pi_{\theta|\phi} : R(\pi_{\theta|\phi} \parallel \pi^*_\theta|\phi) \leq \lambda \right\},$$

$$R(\pi_{\theta|\phi} \parallel \pi^*_\theta|\phi) = \int_\Theta \ln \left( \frac{d\pi_{\theta|\phi}}{d\pi^*_\theta|\phi} \right) d\pi_{\theta|\phi}: \text{KL-distance relative to the benchmark prior.}$$

$\pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_\theta|\phi)$ is abs. continuous with respect to $\pi^*_\theta|\phi$. 
After observing data

The set of posteriors marginalized for $\alpha$:

$$
\Pi^\lambda_{\alpha|X} = \left\{ \pi_{\alpha|X}(\cdot) = \int_\Phi \pi_{\theta|\phi}(\alpha(\theta) \in \cdot) d\pi_{\phi|X} : \pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_\theta|\phi) \forall \phi \right\},
$$
After observing data

The set of posteriors marginalized for $\alpha$:

$$\Pi^\lambda_{\alpha|X} = \left\{ \pi_{\alpha|X}(\cdot) = \int_\Phi \pi_{\theta|\phi}(\alpha(\theta) \in \cdot) d\pi_{\phi|X} : \pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_\theta|\phi) \forall \phi \right\},$$

Report the set of posterior quantities (mean, quantiles, probs): let $h(\alpha) = \alpha, 1\{\alpha \in A\}$, etc.

$$\max_{\pi_{\alpha|X} \in \Pi^\lambda_{\alpha|X}} \int_{\mathbb{R}} h(\alpha) d\pi_{\alpha|X}$$

$$= \max_{\{\pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_\theta|\phi) : \phi \in \Phi\}} \int_\Phi \left( \int_{I\mathcal{S}_\theta(\phi)} h(\alpha(\theta)) d\pi_{\theta|\phi} \right) d\pi_{\phi|X}$$

$$= \int_\Phi \left( \max_{\pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_\theta|\phi)} \int_{I\mathcal{S}_\theta(\phi)} h(\alpha(\theta)) d\pi_{\theta|\phi} \right) d\pi_{\phi|X}$$
Invariance to Marginalization

The KL-neighborhood around $\pi^*_\alpha|\phi$ or around $\pi^*_\theta|\phi$?
Invariance to Marginalization

The KL-neighborhood around $\pi^*_{\alpha|\phi}$ or around $\pi^*_{\theta|\phi}$?

**Marginalization lemma:** Let $\Pi^\lambda(\pi^*_{\alpha|\phi})$ be the KL-neighborhood formed around the $\alpha$-marginal of $\pi^*_{\theta|\phi}$. Then,

$$
\int_{\Phi} \left[ \sup / \inf_{\pi_{\theta|\phi} \in \Pi^\lambda(\pi^*_{\theta|\phi})} \int_{IS_{\theta}(\phi)} h(\alpha(\theta)) d\pi_{\theta|\phi} \right] d\pi_{\phi|X}
$$

$$
= \int_{\Phi} \left[ \sup / \inf_{\pi_{\alpha|\phi} \in \Pi^\lambda(\pi^*_{\alpha|\phi})} \int_{IS_{\alpha}(\phi)} h(\alpha) d\pi_{\alpha|\phi} \right] d\pi_{\phi|X}.
$$

Working with $\Pi^\lambda(\pi^*_{\theta|\phi})$ or $\Pi^\lambda(\pi^*_{\alpha|\phi})$ leads to the same range of posteriors for $\alpha$. 

The range of posteriors

**Theorem:** Suppose $h(\cdot)$ is bounded.

$$
\inf_{\pi_{\alpha|x} \in \Pi_{\alpha|x}^\lambda} E_{\alpha|x}(h(\alpha)) = \int_{\Phi} \left( \int_{\mathcal{I}S_{\alpha}(\phi)} h(\alpha) \, d\pi^\ell_{\alpha|\phi} \right) \, d\pi_{\phi|x}
$$

$$
\sup_{\pi_{\alpha|x} \in \Pi_{\alpha|x}^\lambda} E_{\alpha|x}(h(\alpha)) = \int_{\Phi} \left( \int_{\mathcal{I}S_{\alpha}(\phi)} h(\alpha) \, d\pi^u_{\alpha|\phi} \right) \, d\pi_{\phi|x},
$$

where

$$
d\pi^\ell_{\alpha|\phi} = \frac{\exp(-h(\alpha)/\kappa^\ell_{\lambda}(\phi))}{\int \exp(-h(\alpha)/\kappa^\ell_{\lambda}(\phi)) \, d\pi^*_{\alpha|\phi}} \cdot d\pi^*_{\alpha|\phi}
$$

$$
d\pi^u_{\alpha|\phi} = \frac{\exp(h(\alpha)/\kappa^u_{\lambda}(\phi))}{\int \exp(h(\alpha)/\kappa^u_{\lambda}(\phi)) \, d\pi^*_{\alpha|\phi}} \cdot d\pi^*_{\alpha|\phi},
$$

$\kappa^\ell_{\lambda}(\phi)$ and $\kappa^u_{\lambda}(\phi)$ are the Lagrange multipliers in the constrained optimization problems.
Computing the range of posteriors

**Algorithm:** Assume that $\pi_{\theta|\phi}^*$ or $\pi_{\alpha|\phi}^*$ can be evaluated up to a multiplicative constant

1. Draw many $\phi$’s from $\pi_{\phi|\chi}$

2. At each $\phi$, compute $(\kappa_\lambda^l(\phi), \kappa_\lambda^u(\phi))$ and approximate
   \[ \left[ \int h(\alpha) d\pi_{\alpha|\phi}^l, \int h(\alpha) d\pi_{\alpha|\phi}^u \right] \] using importance sampling (if necessary). Note $(\kappa_\lambda^l(\phi), \kappa_\lambda^u(\phi))$ can be obtained by convex optimizations:
   \[
   \min_{\kappa \geq 0} \left\{ \kappa \ln \int_{\text{IS}_\alpha(\phi)} \exp \left( \pm \frac{\alpha}{\kappa} \right) d\pi_{\alpha|\phi}^* + \lambda \kappa \right\}.
   \]

3. Take the averages of the intervals obtained in Step 2.
Under mild regularity conditions:

- As $n \to \infty$, integration w.r.t. $d\pi_{\phi|\chi}$ can be replaced by plug-in MLE.
- As $\lambda \to \infty$, if $\pi^*_\alpha|\phi$ supports $IS_\alpha(\phi)$, the set of posterior means/quantiles converge to that of Giacomini and Kitagawa (18).
- As $\lambda \to \infty$ and $n \to \infty$, the set of posterior means recovers the true identified set, $IS_\alpha(\phi_0)$. 
**Conditional gamma-minimax**: decision under multiple posteriors

In the empirical application, we compute a robust point estimator for $\alpha$ with the quadratic and absolute losses. As $\lambda \to \infty$ and $n \to \infty$, the gamma-minimax estimator converges to the mid-point of the true identified set.
Conditional gamma-minimax: decision under multiple posteriors

- \( \delta(x) \): decision function, and \( h(\delta(x), \alpha) \): loss function, e.g., \( h(\delta(x), \alpha) = (\delta(x) - \alpha)^2 \).

\[
\min_{\delta(x)} \int \Phi \left[ \max_{\pi_{\alpha|\phi} \in \Pi^\lambda(\pi^*_\alpha|\phi)} \int_{\mathcal{L}_\alpha(\phi)} h(\delta(x), \alpha) d\pi_{\alpha|\phi} \right] d\pi_{\phi|X}
\]

In the empirical application, we compute a robust point estimator for \( \alpha \) with the quadratic and absolute losses.

- As \( \lambda \to \infty \) and \( n \to \infty \), the gamma-minimax estimator converges to the mid-point of the true identified set.
Empirical Example

- BH two-variable SVAR(8), quarterly data for 1970 - 2014.

\[
A_0 \begin{pmatrix} \Delta n_t \\ \Delta w_t \end{pmatrix} = c + \sum_{k=1}^{8} A_k \begin{pmatrix} \Delta n_{t-k} \\ \Delta w_{t-k} \end{pmatrix} + \begin{pmatrix} \epsilon^d_t \\ \epsilon^s_t \end{pmatrix},
\]

\[
A = \begin{pmatrix} 1 & -\beta \\ 1 & -\alpha \end{pmatrix}, \quad (\epsilon^d_t, \epsilon^s_t) \sim \mathcal{N}(0, \text{diag}(d_1, d_2)).
\]
Use BH prior for $\theta$ to generate the single prior $\pi_\phi$ and the benchmark conditional prior $\pi^*_\alpha|\phi$.

The benchmark conditional prior $\pi^*_\alpha|\phi$ is obtained by reparametrizing $\theta \rightarrow (\alpha, \phi)$ and transforming $\pi_\theta$ into $\pi(\alpha, \phi)$. 
Application to SVAR

- BH prior as the benchmark. $\alpha$: labor supply elasticity

![Benchmark posterior (lambda=0)](image)
BH prior as the benchmark. $\alpha$: labor supply elasticity

Range of posteriors: $\lambda = 0.1$

post med bounds

minmax est (abs loss)
Application to SVAR

- BH prior as the benchmark. $\alpha$: labor supply elasticity

Range of posteriors: $\lambda = 0.5$

Giacomini, Kitagawa, Uhlig
Application to SVAR

- BH prior as the benchmark. $\alpha$: labor supply elasticity

-range of posteriors: lambda = 1
- BH prior as the benchmark. $\alpha$: labor supply elasticity

**Range of posteriors: lambda= 2**

- Giacomini, Kitagawa, Uhlig
Application to SVAR

- BH prior as the benchmark. $\alpha$: labor supply elasticity

![Range of posteriors: lambda= 5](image)

Giacomini, Kitagawa, Uhlig
How to elicit $\lambda$?

Given a benchmark prior:

- For each candidate $\lambda$, we can compute the range of priors of the parameter of interest or a parameter for which we have some prior knowledge.
- Compute the ranges of priors for several values of $\lambda$, and pick one that best fits your "vague" prior.
Range of priors (lambda=0.1)

Range of priors (lambda=0.5)

Range of priors (lambda=1)

Range of posteriors (lambda=0.1)

Range of posteriors (lambda=0.5)

Range of posteriors (lambda=1)
1. Why neighborhood around $\pi^*_{\alpha|\phi}$ rather than $\pi^*_\alpha$?
2. Why not multiplier minimax (a la, e.g., Hansen & Sargent)?
3. Is conditional gamma-minimax admissible?
Ho (2019) considers the KL neighborhood around $\pi^*_\alpha$. 

- A prior for $\alpha$ can be updated through the update of $\phi$. Robustness concern should be more serious for the unrevisable part $\pi_{\alpha}\mid\phi$.
- KL neighborhood around $\pi^*_\alpha$ contains multiple priors of $\pi_\phi$, i.e., the marginal likelihood varies over the prior class.
Why neighborhood around $\pi^*_\alpha|\phi$?

Ho (2019) considers the KL neighborhood around $\pi^*_\alpha$.

- A prior for $\alpha$ can be updated through the update of $\phi$. Robustness concern should be more serious for the unrevisable part $\pi^*_{\alpha|\phi}$.
- KL neighborhood around $\pi^*_\alpha$ contains multiple priors of $\pi_\phi$, i.e., the marginal likelihood varies over the prior class.
- Non-unique way of updating the set of priors: Prior-by-prior or Type-II ML updating rule (Good (65, book), Gilboa & Schmeidler (91, JET))?
Why not multiplier minimax?

Along Hansen & Sargent (01), we could formulate the optimal decision under ambiguity in the form of **multiplier minimax**:

\[
\min_{\delta(x)} \int \Phi \left[ \max_{\pi_{\alpha|\phi} \in \Pi^\infty(\pi^*_\alpha|\phi)} \left\{ \int_{S_{\alpha}(\phi)} h(\delta(x), \alpha) d\pi_{\alpha|\phi} - \kappa R(\pi_{\alpha|\phi} \parallel \pi^*_\alpha|\phi) \right\} \right] d\pi_{\phi|x}
\]
Why not multiplier minimax?

- Along Hansen & Sargent (01), we could formulate the optimal decision under ambiguity in the form of **multiplier minimax**:

\[
\min_{\delta(x)} \int \Phi \left[ \max_{\pi_\alpha, \phi \in \Pi^\infty(\pi_{\alpha|\phi})} \left\{ \int I_{S_\alpha}(\phi) h(\delta(x), \alpha) d\pi_\alpha|\phi - \kappa R(\pi_\alpha|\phi\|\pi_{\alpha|\phi}^*) \right\} \right] d\pi_\phi|X
\]

- Given \( \kappa \geq 0 \) and \( \delta \), there exist KL-radius \( \lambda_\kappa(\phi) \) that makes the solution of the gamma minimax equivalent to that of multiplier minimax, but the implied \( \lambda(\phi) \) depends on the loss function \( h(\cdot, \cdot) \).

- We prefer the gamma minimax formulation with fixed \( \lambda \), as we want the set of priors invariant to the loss function.
Unlike unconditional gamma-minimax, conditional gamma-minimax is not guaranteed to be admissible, as the worst-case prior becomes data dependent.

Unconditional gamma-minimax is, on the other hand, difficult to compute, and the difference disappears in large samples.
Related Literatures

- **Bayesian approach to partial identification**: Poirier (97, ET), Moon & Schorfheide (12, Ecta), Kline & Tamer (16, QE), Chen, Christensen, & Tamer (18, Ecta), Giacomini & Kitagawa (18), among others.


- **Decision under Ambiguity**: Schmeidler (89 Ecta), Gilboa & Schmeidler (89 JME), Maccheroni, Marinacci & Rustichini (06 Ecta), Hansen & Sargent (01 AER, 08 book), Strzalecki (11 Ecta), among others
This paper proposes a novel robust-Bayes procedure for a class of set-identified models. It simultaneously deals with the robustness concern in the standard Bayesian method and the excess conservatism in the set-identification analysis.
This paper proposes a novel robust-Bayes procedure for a class of set-identified models. It simultaneously deals with the robustness concern in the standard Bayesian method and the excess conservatism in the set-identification analysis.

Useful for

- **Sensitivity analysis**: perturb the prior of the Bayesian model over the KL-neighborhood, and check how much the posterior result changes.
- **Adding uncertain assumptions** to the set-identified model: Tightening up the identified set by the partially credible benchmark prior.
- **Policy decision** when available data only set-identify a welfare criterion (e.g. treatment choice with observational data).