Discussion of

“Estimation under Ambiguity”

Raffaella Giacomini, Toru Kitagawa, Harald Uhlig

Stéphane Bonhomme
University of Chicago

Conference “Robustness in Economics and Econometrics”

University of Chicago, April 5-6 2019

Classes of priors

• Bayesian estimation hinges on the choice of the prior π.

• In order to assess and mitigate the influence of the prior, "Robust Bayes" approaches consider sets of posterior predictions, for a class of priors $\pi \in \Pi$.

• While conceptually appealing, computing a worst case over a large class Π may give uninformative results.

• Indeed, a motivation for adding prior information in SVARs or DSGEs is that the data alone may not be informative enough.

• This paper develops a sensitivity analysis approach by centering the class Π around a reference prior π^*, and only considering "moderate" deviations from it (such that $KL(\pi||\pi^*) \leq \lambda$).

(θ, ϕ)-modeling

• A key question is whether the prior π is dominated by the data. Partial identification is a leading example where the prior dominates.

• The authors focus on (θ, ϕ) models, where ϕ is point-identified, whereas θ is independent of the data given ϕ. So π_ϕ is dominated, and $\pi_{\theta|\phi}$ is not.

• This parameterization is insightful, and it is also natural in examples such as supply and demand or SVARs.

• It seems essential to the authors' approach.

• However, the distinction between ϕ and θ may not always be obvious. Moreover, it does not allow for "weak" identification scenarios, where the prior may still play important role.

θ-modeling

• Is it possible to modify the approach so that it applies to any parameter θ, without presuming the nature of identification of its components?

• For this, consider an unconditional Gamma-minimax approach (under squared loss for concreteness):

$$\inf_{\delta} \mathcal{R}^{\lambda}(\delta) = \inf_{\delta} \sup_{\pi \in \Pi^{\lambda}(\pi^*)} \mathbb{E}_{\pi}[(\alpha(\theta) - \delta(X))^2].$$

• $\mathcal{R}^{\lambda}(\delta)$ coincides with frequentist minimax risk when the class Π of priors is unrestricted.

• In general, this differs from the conditional approach based on posterior risk that the authors consider, since here the expectation is not conditional on the sample $X_1, ..., X_n$. The two approaches agree when θ and X are independent.

Small-λ approximation

• To make progress, I rely on a small-λ approximation ("local" robustness).

• For small λ, the risk can be expanded as:

$$\mathcal{R}^{\lambda}(\delta) \approx \underbrace{\mathbb{E}_{\pi^*}[(\alpha(\theta) - \delta(X))^2]}_{\text{Bayes risk}} + \underbrace{\lambda^{\frac{1}{2}} \left\{ \mathrm{Var}_{\pi^*} \left(\mathbb{E}[(\alpha(\theta) - \delta(X))^2 \,|\, \theta] \right) \right\}^{\frac{1}{2}}}_{\text{Robustness adjustment}}.$$

• Here the Bayesian risk is adjusted for the possibility that π differs from π^*.

• This expression allows for partial or irregular identification of θ.

Small-λ approximation (cont.)

- Recall the expansion:

$$\mathcal{R}^{\lambda}(\delta) \approx \underbrace{\mathbb{E}_{\pi^*}[(\alpha(\theta)-\delta(X))^2]}_{\text{Bayes risk}} + \underbrace{\lambda^{\frac{1}{2}}\left\{\mathsf{Var}_{\pi^*}\left(\mathbb{E}[(\alpha(\theta)-\delta(X))^2 \mid \theta]\right)\right\}^{\frac{1}{2}}}_{\text{Robustness adjustment}}.$$

- Unlike posterior risk, minimizing $\mathcal{R}^{\lambda}(\delta)$ is a functional optimization problem (albeit with a convex objective function).

- I have not been able to compute a similar small-λ approximation of posterior risk.

- However, a fixed-λ characterization of $\mathcal{R}^{\lambda}(\delta)$ is also available. Is risk minimization computationally feasible in that case?

Comparison to frequentist estimation

• Suppose now θ is a latent variable, π is its unknown distribution, and we wish to estimate $\int \alpha(\theta)\pi(\theta)d\theta$ for some function α.

• This type of problems arises in panel data models, where one wishes to estimate an average effect.

• The frequentist minimax risk is then:

$$\mathcal{R}_f^{\lambda}(\delta) = \sup_{\pi \in \Pi^{\lambda}(\pi^*)} \mathbb{E}_{\pi}\left[\left(\int \alpha(\theta)\pi(\theta)d\theta - \delta(X)\right)^2\right].$$

• We see that:

$$\mathcal{R}_f^{\lambda}(\delta) = \sup_{\pi \in \Pi^{\lambda}(\pi^*)} \left\{\mathbb{E}_{\pi}\left[(\alpha(\theta) - \delta(X))^2\right] - \mathsf{Var}_{\pi}\left(\alpha(\theta)\right)\right\}.$$

So this measure of frequentist risk is closely related to the Bayesian risk the authors consider.

Priors or models?

- Martin and I have characterized the form of the δ functions that minimize $\mathcal{R}_f^{\lambda}(\delta)$, under a small-$\lambda$ approximation.

- The optimal estimator can be computed by solving a linear system of functional equations.

- The frequentist approach is useful, since it allows us to compute asymptotically valid confidence intervals.

- The link between the two approaches comes from the dual role of $\pi(\theta)$ when θ is a latent variable: as the distribution of θ, and as a prior for θ_i, for $i = 1, ..., n$.

- This example raises the question: should we think of π as a "prior" or as part of the "model"?

- Adding (economic) structure to π^* and $\Pi^{\lambda}(\pi^*)$ could help make sensitivity analysis more interpretable.