Discussion of "Adapting to Misspecification" by Armstrong, Kline and Sun

Discussant: Xu Cheng

Northwestern Interactions Workshop
September 22, 2023

- A fast growing literature on estimation/inference under model misspecification
  - consider perturbations of a correctly specified model
  - sensitivity analysis and robust estimation/inference e.g., Chen, Tamer, Torgovitsky (2011), Andrews, Gentzkow, Shapiro (2017), Armstrong and Kolesar (2018, 2021), Bonhomme and Weidner (2021), Christensen and Connault (2023), etc
  - robustness/efficiency under misspecified models
    e.g., Bugni, Canay, Guggenberger (2012), Hansen (2016), Cheng, Liao, Shi (2019), Fessler and Kasy (2019), de Chaisemartin and D'Haultfœuille (2020), etc

- This paper makes an important contribution on robustness and efficiency tradeoff with misspecified models
  - there exists an unrestricted Y_U estimator that is asy unbiased, e.g., valid instruments
  - an restricted estimator Y_R with asy bias $b \in B$ e.g., add additional invalid instruments, $b = \sqrt{n}\mathbb{E}[ZU]$
  - misspecification brings in bias but reduces variance
  - b and B are both unknown

- Some existing results in similar setups and challenges
  - pre-test / post model selection estimator is bad (Leeb and Pötscher, 2005)
  - various data-dependent smooth average of Y_U and Y_R
    - Hansen (2016), Cheng, Liao, Shi (2019), Fessler and Kasy (2019)
    - e.g., Cheng, Liao, Shi (2019) derive the risk of the averaging estimator as a function of b and plug in its unbiased estimator
    - the key is to show uniform dominance – uniformly over $b \in [0, \infty)$, the averaging estimator always has smaller risk than the unrestricted estimator Y_U, for a vector of parameters
    - however, this James-Stein type shrinkage phenomenon does not work for a scalar parameter as in the present paper
- This paper studies a scalar parameter and the minimax risk

- The main challenge is $b \in B$ and the upper bound B is unknown
  - a creative solution based on adaptation regret: the price to pay without knowing B
  - if we know B, we can construct an estimator with min worst case risk $R^*(B)$
  - if we don't know B, we obtain worst case risk $R_{max}(B, \delta)$ for the estimator δ
  - choose δ to minimize $\sup_B \frac{R_{max}(B,\delta)}{R^*(B)}$
  - near optimal by a multiplicative factor
  - convert to minimax estimation with scaled loss for easy computation
- Get back to comparison with Y_U, the paper has a very nice result on adaptation with a worst case risk upper bound

- Some other interesting questions that I have got on a setup with Y_U and Y_R
  - How about averaging more than two estimators?
  - The paper has an extension to multiple restricted estimators!
  - What if the baseline model is also misspecified, maybe to a less degree?
  - What if the baseline model does not provide sufficient identification?

Summary

- The paper provides a great solution to the robustness and efficiency trade off
- focus on the challenging case of a scalar parameter
- introduce the idea of adaptive estimation to allow for unknown bound on the degree of misspecification
- sophisticated computation method
- empirical applications in a wide range of scenarios