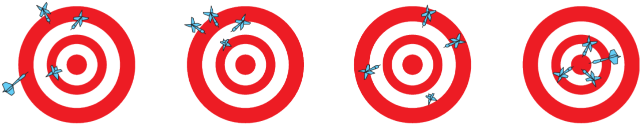
Bias and Variance of Estimators for Hellinger Speedless Points

David Pacini
University of Bristol

ESEM 2023
Barcelona, 28 August 2023

Big Picture

Research: Nonexistence finite-variance unbiased estimators



A B C D

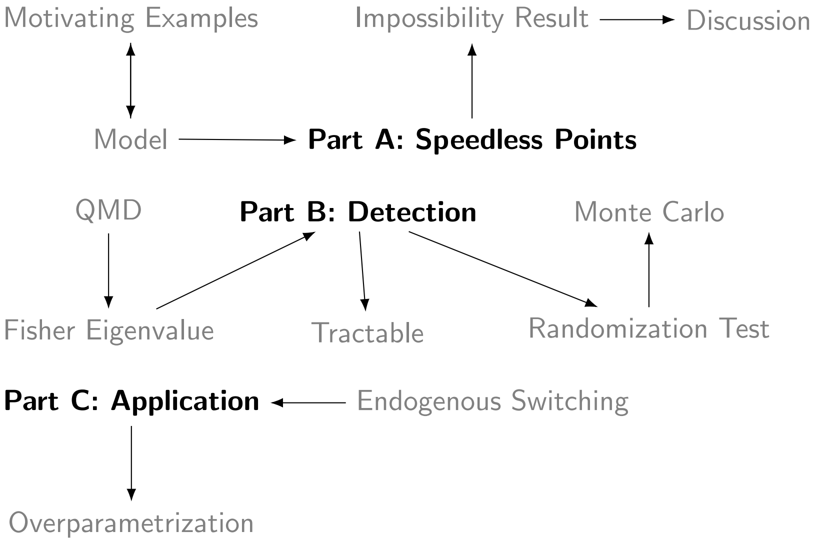
Articulate: bias correction, evaluating criteria,
normal approximations, algorithm/computation

Context: estimation with endogeneity, self-selection, switching...

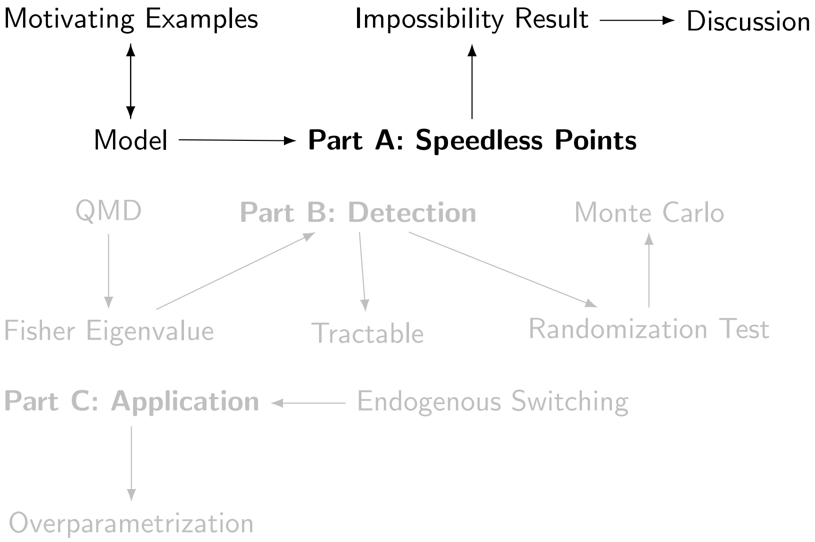
Finite-Variance Unbiased Estimators

- Sufficient for nonexistence: Hellinger speedless point
- Detection via least eigenvalue
- Overparametrization and nonexistence

Project Overview



This Presentation



Motivating Examples

Normal parametric density featuring

- Endogeneity as in linear IV

$$Y_{1i} = x'_{1i}\beta_1 + Y_{2i}\beta_2 + \epsilon_{1i}$$
$$Y_{2i} = x'_{1i}\gamma_1 + x'_{2i}\gamma_2 + \epsilon_{2i}$$

- Self-selection as in Gronau (1974, JPE)

$$Y_{1i} = Y^{\star}_{1i} Y_{2i}, \quad Y^{\star}_{1i} = x'_{1i}\beta_1 + \epsilon_{1i};$$
$$Y_{2i} = 1(x'_{1i}\gamma_1 + x'_{2i}\gamma_2 - \epsilon_{2i} \geq 0)$$

- Switching as in Lokshin and Ginskaya (2009, WBER)

$$Y_{1i} = 1(x'_{1i}\beta_1 + Y_{2i}x'_{1i}\beta_2 - \epsilon_{1i} \geq 0)$$
$$Y_{2i} = 1(x'_{1i}\gamma_1 + x'_{2i}\gamma_2 - \epsilon_{2i} \geq 0)$$

- Use common framework for nonexistence

Statistical Model

- IID sample $\{Y_i\}_{i=1}^N$, covariates as fixed
- Parametric density L_2-family:

$$\mathcal{F}_\Theta := \left\{ f_\theta : \mathcal{Y} \to \mathbb{R},\ f_\theta \geq 0,\ \int f_\theta = 1,\ f_\theta < \bar{g},\ \int \bar{g}^2 < \infty,\ \theta \in \Theta \right\}$$

- Unbiased estimator

$$\int \hat{\theta}_N f_\theta^N - \theta = 0 \text{ for every } \theta \in \Theta$$

Part A: Hellinger Speedless Point

- Bullseye in parameter space: $\theta_\star$ Bullseye
- Smooth curve through bullseye: Curve

$$[0,1] \ni \epsilon \mapsto c(\epsilon) \in \Theta \text{ such that } c(0) = \theta_\star$$

- Hellinger distance to bullseye: Hellinger

$$h(c(\epsilon), \theta_\star) := \sqrt{\int \frac{\left| f_{c(\epsilon)}^{1/2} - f_{\theta_\star}^{1/2} \right|^2}{2} d\mu}$$

- Speed at bullseye: $\lim_{\epsilon \to 0} \frac{h(c(\epsilon), \theta_\star)}{|\epsilon|}$
- Bullseye is speedless if:

$$\inf_c \lim_{\epsilon \to 0} \frac{h(c(\epsilon), \theta_\star)}{|\epsilon|} = 0$$

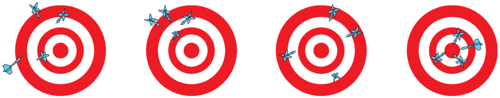
Impossibility Result

Proposition:

Set bullseye to $\theta_\star = (\beta_\star, \gamma_\star)$

If bullseye is speedless, $\beta_\star$, $\gamma_\star$, or both do not have a

finite-variance unbiased



A B C D

Discussion

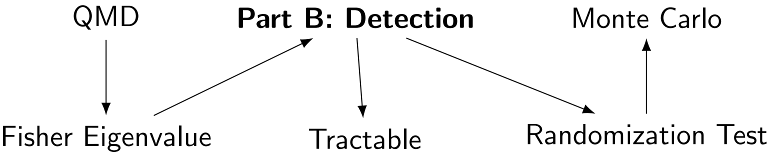
Bullseye speedless $\implies$ nonexistence finite-variance unbiased

- gist: Cauchy-Schwarz inequality
- $\beta_\star$, $\gamma_\star$, or both?
- nonidentifiable bullseye $\implies$ speedless, but not viceversa
- examples: linear triangular IV, self-selection, squared location:

$$\Theta = \mathbb{R}, \theta_\star = 0, f_\theta(y) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(y-\theta^2)^2}{2}\right)$$

- Speedless are Lebesgue measure zero

Detecting Speedless Points



Proposition:

$$\underbrace{\inf_{c} \lim_{\epsilon \to 0} \frac{h(c(\epsilon), \theta_\star)}{|\epsilon|}}_{Hellinger} = \min_{q : q^\top q = 1} q^\top \underbrace{\mathbb{E}_{\theta_\star}(s_{\theta_\star} s_{\theta_\star}^\top)}_{Fisher} q$$

Calculating least eigenvalue:

- QR
- Jacobi
- Rayleigh

Application

Part C: Application ← Endogenous Switching

↓

Overparametrization

Endogenous switching model:

$$Y_{1i} = 1(x_{1i}'\beta_1 + Y_{2i}x_{1i}'\beta_2 - \epsilon_{1i} \geq 0)$$
$$Y_{2i} = 1(x_{1i}'\gamma_1 + x_{2i}'\gamma_2 - \epsilon_{2i} \geq 0)$$

- No endogeneity: $\rho_\star = 0$
- No switching: $\beta_{2,\star} = 0$
- No instruments: $\gamma_{2,\star} = 0$

Illustrating Use

- Migration and labor market participation of couples

  Bivariate binary endogenous switching model

- Pull out:

  Nonexistence finite-variance unbiased estimators

  Complete bias-correction not desirable

  If maximum likelihood, first-order algorithm

  Normal large-sample inference unreliable

Positioning

My paper: speedless + finite-sample + detection test

- Nonidentifiable points: speedless? test?
- Vertical asymptotae reduced-to-structural mapping: speedless? test?
- Fisher information regular paths: finite-sample? test?
- Hellinger modulus continuity: speedless? test?

References:

Metrics: Sargan (1983); Gourieroux and Monfort ('95, Chapter 5), Hirano and Porter ('15, ER), Khan and Nepekilov ('18, QE);
Stats: Halmos ('46, AMS), Pitman ('76, Chapter 1), van der Vaart ('91, AoS), Brown and Liu ('93, AoS)

Conclusion

Hellinger speedless points make finite-variance unbiased estimation impossible

- Detection direct or via randomization test
- Consequences

  some speedless are identifiable

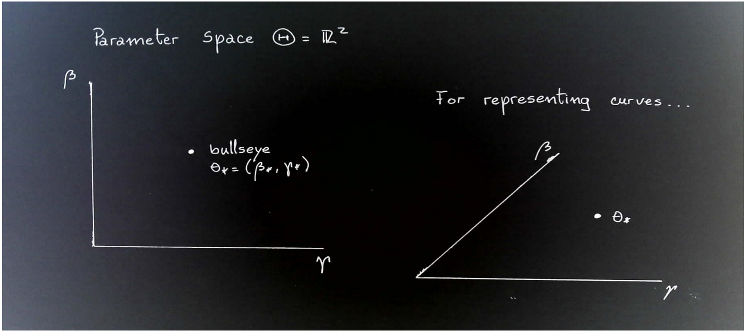
  overparametrization and bias correction

  alternative evaluation criteria

  computation and normal inference

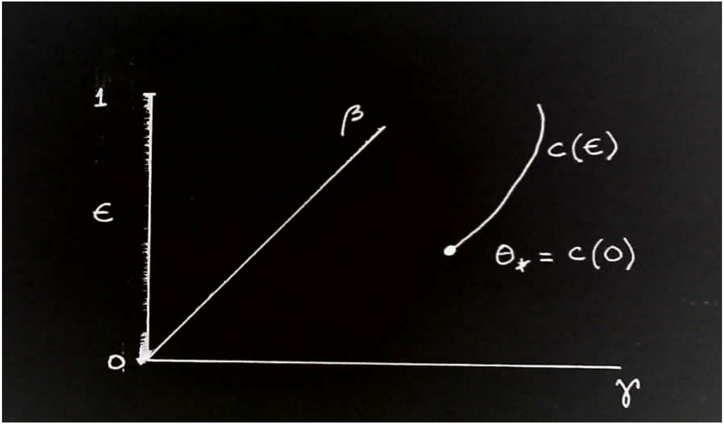
Bullseye in Parameter Space

Back



Curve Passing through Bullseye

Back



Hellinger Distance to Bullseye

Back