Optimal Confidence Procedures for Partially Identified Parameters

Ian Pitman

University of Chicago

November 8, 2026

Motivation

- Partially identified models are widespread in empirical economics
  - **Advantage:** enable relaxation of implausible assumptions
  - **Drawback:** complicate statistical inference
- Imbens and Manski (2004) and ensuing work emphasize coverage properties
- Little is known about length properties
- Under point identification, conventional CIs $\left[\hat{\theta} \pm 1.96\frac{\sigma}{\sqrt{n}}\right]$ are length optimal

Theoretical Question

How can we extend classical optimality theory to partially identified models?

Quantitative Question

How far from optimal are existing confidence procedures?

Outline of the Talk

Model and Examples

Statistical Decision Problem

Reduction to an Optimal Testing Problem

Duality and Computation

Numerical Results

Running Example: Rambachan and Roth (2023)

Setup and Notation

Balanced panel: $\{Y_{it}, D_i\}_{i=1,\, t=-\underline{T}}^{n,\ \overline{T}}$

Potential outcomes: $Y_{it} = (1 - D_i)Y_{it}(0) + D_i Y_{it}(1)$

Treatment timing: $t = 1$ (absorbing, non-staggered)

No anticipation: $Y_{it}(0) = Y_{it}(1)$ for $t < 1$

Difference-in-Differences estimand:

Example Continued

$$\begin{aligned}\mu_t &\equiv \big(\mathbb{E}[Y_{it} \mid D_i = 1] - \mathbb{E}[Y_{it} \mid D_i = 0]\big) - \big(\mathbb{E}[Y_{i0} \mid D_i = 1] - \mathbb{E}[Y_{i0} \mid D_i = 0]\big) \\ &= \underbrace{\mathbb{E}[Y_{it}(1) - Y_{it}(0) \mid D_i = 1]}_{\text{ATT in period } t} + \underbrace{\big(\mathbb{E}[Y_{it}(0) - Y_{i0}(0) \mid D_i = 1] - \mathbb{E}[Y_{it}(0) - Y_{i0}(0) \mid D_i = 0]\big)}_{\text{Difference in trends in untreated potential outcomes from } 0 \text{ to } t} \\ &\equiv \tau_t + \delta_t\end{aligned}$$

Relaxing the Parallel Trends Assumption

Stacking pre-period (excluding 0) and post-period equations, since $\tau^{\text{pre}} = 0$,

$$\underbrace{\begin{bmatrix} \mu^{\text{post}} \\ \mu^{\text{pre}} \end{bmatrix}}_{\equiv \mu} = \underbrace{\begin{bmatrix} \mathbf{I}_{\overline{T}} \\ \mathbf{0}_{\underline{T} \times \overline{T}} \end{bmatrix}}_{\equiv K} \tau^{\text{post}} + \underbrace{\begin{bmatrix} \delta^{\text{post}} \\ \delta^{\text{pre}} \end{bmatrix}}_{\equiv \delta}$$

Model

$$\mu = K\tau^{\text{post}} + \delta$$
$$\theta = l'\tau^{\text{post}}$$

for some $(\tau^{\text{post}}, \delta) \in \mathbb{R}^{k_\tau} \times \Delta$

Experiment

$$\hat{\mu}_n \sim \mathcal{N}(\mu, \Omega/n)$$

- **Parallel trends:** $\Delta = \{0\} \implies \theta = l'\tau^{\text{post}} = l'\mu^{\text{post}}$ point identified
- Building on Manski and Pepper (2018), Rambachan and Roth (2023) propose relaxing parallel trends by expanding Δ

Examples of Nuisance Parameter Sets

Relative Magnitudes

$$\Delta^{\mathsf{RM}}(\bar{M}) \equiv \left\{ \delta : |\delta_{t+1} - \delta_t| \leq \bar{M} \max_{s<0} |\delta_{s+1} - \delta_s| \text{ for all } t \geq 0 \right\}$$

Relative Magnitudes with a Decreasing Differential Trend

$$\Delta^{\mathsf{RMD}}(\bar{M}) \equiv \Delta^{\mathsf{RM}}(\bar{M}) \cap \left\{ \delta : \delta_{t+1} - \delta_t \leq 0 \text{ for all } t \right\}$$

Application

A Flexible Class of Models

Model

$$\mu = K\tau + \delta \qquad \theta = l'\tau \qquad \text{for some} \quad (\tau, \delta) \in \mathbb{R}^{k_\tau} \times \Delta$$

Experiment

$$X \sim \mathcal{N}(\mu, \Sigma) \equiv P_\mu$$

- K, l, Δ, and Σ are all **known**
- Moments $\mu \in \text{range}(K) + \Delta \equiv \mathcal{M}$ are consistent with the model
- Armstrong and Kolesár (2018, 2021) establish that Donoho (1994)'s CI has near minimax expected length when Δ convex, centrosymmetric ($\delta \in \Delta \iff -\delta \in \Delta$)
- **Flexibility:** I leave the specification of Δ *entirely* up to the researcher's discretion

Union Bounds and Intersection Bounds

- $\Delta^{\text{RM}}(\bar{M})$ is centrosymmetric, but *not* convex
- Under $\Delta^{\text{RM}}(\bar{M})$, the identified set for $\theta = \tau_t$ takes the form of union bounds

$$\Theta(\mu) = \left[\min_{s<0} \left\{ \mu_t - t\bar{M}|\mu_{s+1} - \mu_s| \right\},\ \max_{s<0} \left\{ \mu_t + t\bar{M}|\mu_{s+1} - \mu_s| \right\} \right]$$

- Union bounds more generally also fit into my framework
  - Often appear in sensitivity analysis (Bei, 2025)
- Intersection bounds (swapping the $\max$ and $\min$) are also compatible
  - Often appear in program evaluation (Manski, 1990)

Several Applications Compatible with my Framework

Treatment Effects and Average Counterfactuals

- Lee (2009), Gundersen et al. (2012)
- de Quidt et al. (2018), Kolesár et al. (2026)

Sign Restrictions and Bounds on OVB

- Li and Müller (2021), Ketz and McCloskey (2025), Cox (2026)

Discrete Choice

- Katz (2007), Pakes et al. (2015), Canay et al. (2023)

Sensitivity Analysis

- Kolesár and Rothe (2018), Rambachan and Roth (2023)
- Nevo and Rosen (2012), Masten and Poirier (2021)

Outline of the Talk

Model and Examples

Statistical Decision Problem

Reduction to an Optimal Testing Problem

Duality and Computation

Numerical Results

Background on Statistical Decision Theory

- **Decision theory:** a unifying framework for defining "optimality" (Wald, 1950)
  - estimation, hypothesis testing, treatment assignment, experimental design, etc.
- I take a decision-theoretic perspective on confidence procedures

Essential Elements of a Statistical Decision Problem

Experiment:	Observe $X \sim P_\mu$ for some $\mu \in \mathcal{M}$
Decision Rule:	What confidence set to report based on X
Loss Function:	How costly it is to report a given set under P_μ
Risk Function:	Expected loss for a given confidence procedure under P_μ
Risk Aggregation:	How to trade off risk across states of the world μ

Action Space and Confidence Procedure

- The researcher takes an action by reporting a confidence set
- **Randomization device:** $U \sim \text{Unif}[0,1]$ independent of X
- **Confidence procedure:** measurable map from sample space $\to$ action space

$$S : \mathbb{R}^{k_\mu} \times [0,1] \to \mathcal{A}$$

- Action space $\mathcal{A}$ is the collection of all measurable sets Technical Definition
  - Two sets are considered equivalent if they only differ on a set of measure zero
- The researcher need not report an interval unless it is optimal to do so

Class of Confidence Procedures with Correct Coverage

- For $\alpha \in (0, 0.5]$, I require $1 - \alpha$ coverage of the partially identified parameter

Imbens and Manski (2004) Coverage Requirement

$$S \in \mathcal{S}_\alpha \equiv \left\{ S \in \mathcal{S} : \inf_{\mu \in \mathcal{M}} \inf_{t \in \Theta(\mu)} \Pr_{P_{\mu,U}}(t \in S(X, U)) \geq 1 - \alpha \right\}$$

Technical Definition

Decision Theory Checklist

Experiment ✓

Decision Rule ✓

Loss Function

Risk Function

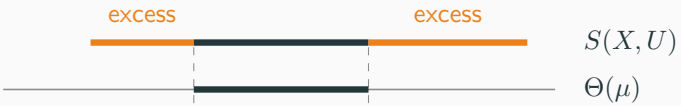
Risk Aggregation

Loss and Risk Functions

Excess Length Loss and Risk

Loss: $L : \mathcal{M} \times \mathcal{A} \to \overline{\mathbb{R}}_+ \qquad L(\mu, A) \equiv \lambda\big(A \setminus \Theta(\mu)\big)$

Risk: $R : \mathcal{M} \times \mathcal{S} \to \overline{\mathbb{R}}_+ \qquad R(\mu, S) \equiv \mathbb{E}_{P_{\mu,U}}\big[L(\mu, S(X, U))\big]$



Interpretation: not costly to report values of θ consistent with the model

Decision Theory Checklist

Experiment ✓

Decision Rule ✓

Loss Function ✓

Risk Function ✓

Risk Aggregation

Aggregating Risk

- To fix ideas, I focus on **minimax** risk
  - Conventional two-sided CIs are minimax expected length optimal (Stein, 1962)
  - I can also accommodate minimax regret and Bayes risk
- $S^\star \in \mathcal{S}_\alpha$ is minimax optimal if it solves

$$\inf_{S \in \mathcal{S}_\alpha} \sup_{\mu \in \mathcal{M}} R(\mu, S)$$

- **Obstacle:** the choice variable is a *function* mapping (X, U) to sets

Question

How can we leverage the model's structure to make the problem tractable?

Outline of the Talk

Model and Examples

Statistical Decision Problem

Reduction to an Optimal Testing Problem

Duality and Computation

Numerical Results

Background on Test Inversion

- **Test inversion:** main method of forming confidence sets in PID models
- Exploits duality between confidence sets and tests

  95% confidence set for $\theta = \{t \in \mathbb{R} : H_0 : t \in \Theta(\mu)$ *not* rejected at the 5% level$\}$

- Observe noisy estimate of μ, not θ, so helpful to write

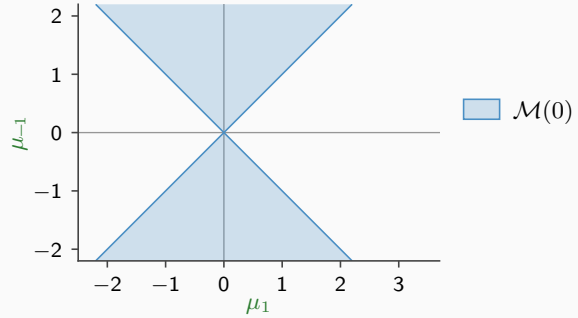
$$t \in \Theta(\mu) \iff \mu \in \mathcal{M}(t)$$

- How can we simplify the problem of testing $H_0 : \mu \in \mathcal{M}(t)$?

Example Continued: Rambachan and Roth (2023)

Three-period DiD model

$$\begin{bmatrix}\mu_1\\ \mu_{-1}\end{bmatrix} = \begin{bmatrix}1\\ 0\end{bmatrix}\tau_1 + \begin{bmatrix}\delta_1\\ \delta_{-1}\end{bmatrix}, \quad \Delta^{\text{RM}}(1) = \left\{ \begin{bmatrix}\delta_1\\ \delta_{-1}\end{bmatrix} : |\delta_1| \leq |\delta_{-1}| \right\}, \quad \Theta(\mu) = \left[\mu_1 \pm |\mu_{-1}|\right]$$

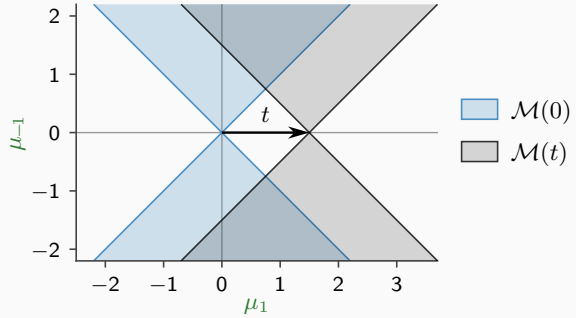


Setup

Example Continued: Rambachan and Roth (2023)

Three-period DiD model

$$\begin{bmatrix}\mu_1\\ \mu_{-1}\end{bmatrix} = \begin{bmatrix}1\\ 0\end{bmatrix}\tau_1 + \begin{bmatrix}\delta_1\\ \delta_{-1}\end{bmatrix}, \quad \Delta^{\text{RM}}(1) = \left\{ \begin{bmatrix}\delta_1\\ \delta_{-1}\end{bmatrix} : |\delta_1| \leq |\delta_{-1}| \right\}, \quad \Theta(\mu) = \left[\mu_1 \pm |\mu_{-1}|\right]$$



Setup

Recentering the Moments

- The model has a linear and additively separable structure

$$\begin{aligned} \mu &= K\tau + \delta \\ \theta &= l'\tau \end{aligned} \quad \text{for some} \quad (\tau, \delta) \in \mathbb{R}^{k_\tau} \times \Delta$$

- Null set **translates** with t, that is, $\mathcal{M}(t) = tKl + \mathcal{M}(0)$, so

$$\mu \in \mathcal{M}(t) \iff \mu - tKl \in \mathcal{M}(0)$$

- Suffices to test whether **recentered** moments consistent with $\theta = 0$

Projecting Out Uninformative Directions

- Some directions in μ space may be uninformative about θ

$$V \equiv \left\{ v \in \mathbb{R}^{k_\mu} : tv + \mathcal{M}(0) = \mathcal{M}(0) \text{ for all } t \in \mathbb{R} \right\}$$

- I **project** them out to reduce the dimension of the problem

$$\mu - tKl \in \mathcal{M}(0) \iff Q(\mu - tKl) \in Q\mathcal{M}(0)$$

- The rows of Q form an orthogonal basis of $V^\perp$, chosen so that $QKl = e_1$

Assumption 1: Informativeness

$Kl \notin V$.

The Reduced Testing Problem

- Putting everything together,

$$t \in \Theta(\mu) \iff Q(\mu - tKl) \in Q\mathcal{M}(0)$$

Reduced Testing Problem

Reduced Moments: $\nu \equiv Q(\mu - tKl) \in Q\mathcal{M}$

Reduced Statistic: $Y \equiv Q(X - tKl) \sim \mathcal{N}(\nu, Q\Sigma Q') \equiv P_\nu$

Null Hypothesis: $H_0 : \nu \in Q\mathcal{M}(0)$

Level-α Tests: $\Phi_\alpha \equiv \left\{\varphi : \mathbb{R}^{k_\nu} \to [0,1] \text{ s.t. } \sup_{\nu \in Q\mathcal{M}(0)} \mathbb{E}_{P_\nu}[\varphi(Y)] \leq \alpha\right\}$

- Only the **first** component of ν and Y depends on t

$$\nu = \begin{bmatrix} \nu_1 \\ \nu_{-1} \end{bmatrix} = \begin{bmatrix} Q_1\mu - t \\ Q_{-1}\mu \end{bmatrix}$$

Extending the Ghosh–Pratt Identity

- **Conventional wisdom:** inverting a family of tests with good power properties $\implies$ confidence procedure with good length properties
- The Ghosh–Pratt identity makes this precise (Ghosh, 1961; Pratt, 1961)

Ghosh–Pratt Identity under Partial Identification

Expected excess length = integrated type II error probability

$$R(\mu, S) = \int_{\mathbb{R}\setminus\Theta(\mu)} \Pr_{P_{\mu,U}}(t \in S(X,U))\,\mathrm{d}t$$

$$R(\nu_{-1}, \varphi) \equiv \int \underbrace{\mathbb{1}\{\nu \notin Q\mathcal{M}(0)\}}_{\text{indicator for the alternative}} \overbrace{\mathbb{E}_{P_\nu}[1-\varphi(Y)]}^{\text{type II error}}\,\mathrm{d}\nu_1$$

Additional Regularity Assumptions

Assumption 2: Normalization

$l \in \mathbb{R}^{k_\tau}$ is a nonzero vector normalized so that $\|l\| = 1$.

Assumption 3: Nuisance Parameter Set

$\Delta \subseteq \mathbb{R}^{k_\mu}$ is non-empty and σ-compact.

Assumption 4: Positive Definiteness

The covariance matrix Σ is positive definite, $\Sigma \succ 0$.

The Main Result

Theorem 1: Reduction to an Optimal Testing Problem

Under Assumptions 1 to 4, the optimal testing problem

$$\inf_{\varphi \in \Phi_\alpha} \sup_{\nu_{-1} \in Q_{-1}\mathcal{M}} R(\nu_{-1}, \varphi)$$

admits a solution $\varphi^\star$. Moreover, the test inversion confidence procedure

$$S_{\varphi^\star}(x, u) \equiv \{t \in \mathbb{R} : u \leq 1 - \varphi^\star(Q(x - tKl))\}$$

has minimax expected excess length, meaning it solves

$$\inf_{S \in \mathcal{S}_\alpha} \sup_{\mu \in \mathcal{M}} R(\mu, S).$$

Takeaways from the Main Result

- Performance of test inversion CS depend on the **joint** behavior of a family of tests
- Theorem 1 decouples the problem across candidate values $t \in \mathbb{R}$
- The researcher must only solve a **single** optimal testing problem
- Compared to conventional test inversion procedures,
  - Shape of $\mathcal{M}(t)$ is immaterial; what matters is how it *translates* with t
  - Computational bottleneck is the testing problem, *not* recomputing critical values

Outline of the Talk

Model and Examples

Statistical Decision Problem

Reduction to an Optimal Testing Problem

Duality and Computation

Numerical Results

Solving the Optimal Testing Problem

- Applying Theorem 1 requires solving the minimax optimal testing problem

$$\inf_{\varphi \in \Phi_\alpha} \sup_{\nu_{-1} \in Q_{-1}\mathcal{M}} R(\nu_{-1}, \varphi)$$

- **Challenge:** searching over tests $\varphi \in \Phi_\alpha$
- Duality theory enables us to **only** search over tests of Neyman–Pearson form (Krafft and Witting, 1967)

The Dual Problem

- Let $\pi \in \Pi$ denote weights over the risk parameters $Q_{-1}\mathcal{M}$, and $\Gamma \in \mathcal{G}$ denote weights over the null $Q\mathcal{M}(0)$
- Then for $\varphi_{\pi,\Gamma}$ a test of Neyman–Pearson form, NP

$$\sup_{\pi\in\Pi}\sup_{\Gamma\in\mathcal{G}} \int R(\nu_{-1}, \varphi_{\pi,\Gamma})\, \pi(\mathrm{d}\nu_{-1}) + \int \left(\mathbb{E}_{P_\nu}[\varphi_{\pi,\Gamma}(Y)] - \alpha\right) \Gamma(\mathrm{d}\nu)$$

  is the dual of the optimal testing problem.
- If $(\pi^\star, \Gamma^\star)$ solve the dual, then $\varphi_{\pi^\star,\Gamma^\star}$ solves the primal
- Aradillas Fernández et al. (2025) give **inner**-problem guarantees for finite $Q\mathcal{M}(0)$
- I extend their approach to solve the **outer** problem simultaneously SMA

Outline of the Talk

Model and Examples

Statistical Decision Problem

Reduction to an Optimal Testing Problem

Duality and Computation

Numerical Results

Calibrated Simulation

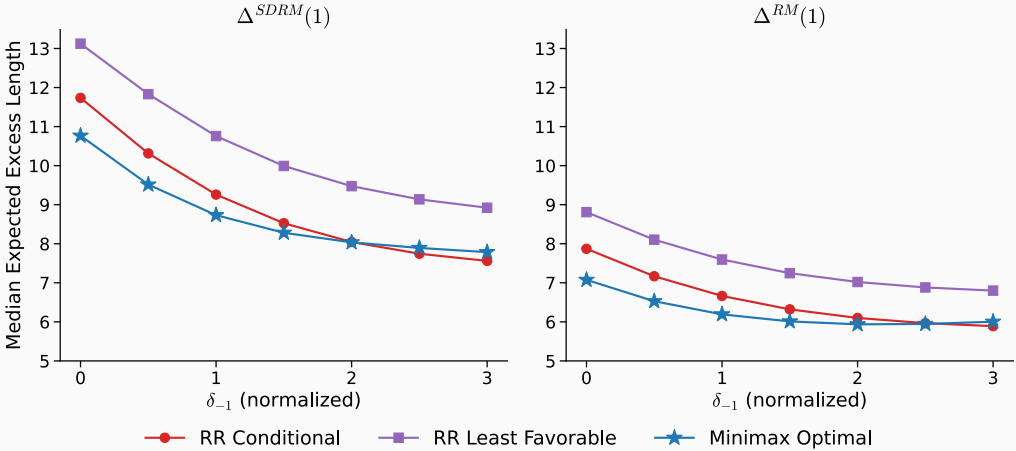
- I adapt the simulation design of Rambachan and Roth (2023) to benchmark their recommended procedures against mine

Simulation Design

Covariance Matrices:	Calibrated from six papers with $\underline{T} \leq 4$
Target Parameter:	$\theta = \tau_1$
True Parameters:	$\tau = 0 \implies \mu = \delta$, with $\delta_t = 0$ for all $t \neq -1$
"Pulse" Pre-Trend:	δ_{-1} from 0 to 3 standard deviations
Benchmarks:	Test inversion based on Andrews et al. (2023)

Definition

Calibrated Simulation Results



Reduction in Risk of Optimal Procedure at $\delta_{-1} = 0$ Across Papers

	RR Conditional		RR Least Favorable	
Restriction	Median	Range	Median	Range
$\Delta^{\text{SDRM}}(1)$	7.9%	2.5% – 9.5%	18.4%	15.5% – 20.1%
$\Delta^{\text{RM}}(1)$	10.9%	3.9% – 11.6%	20.2%	16.0% – 21.3%

Application: Event Study from Iovino et al. (2026)

- **Question:** how did the 2018 corporate tax cut affect CO_2-intensive firms?
- **Data:** US public firms, 2015–2021; outcome log capital

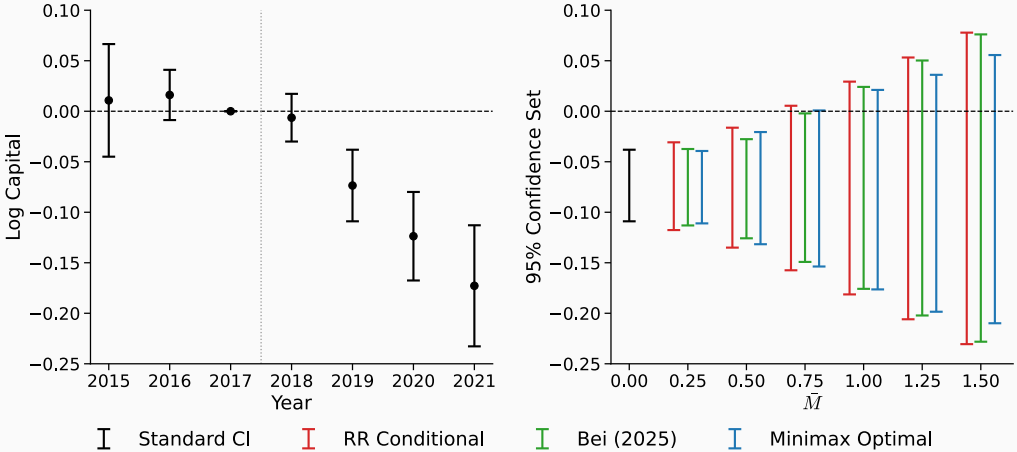
Event Study Specification

$$Y_{it} = \sum_{s \neq 2017} \mu_s \, \mathbb{1}\{t = s\} \times \text{Intensive}_i + \alpha_i + \gamma_t + \varepsilon_{it}$$

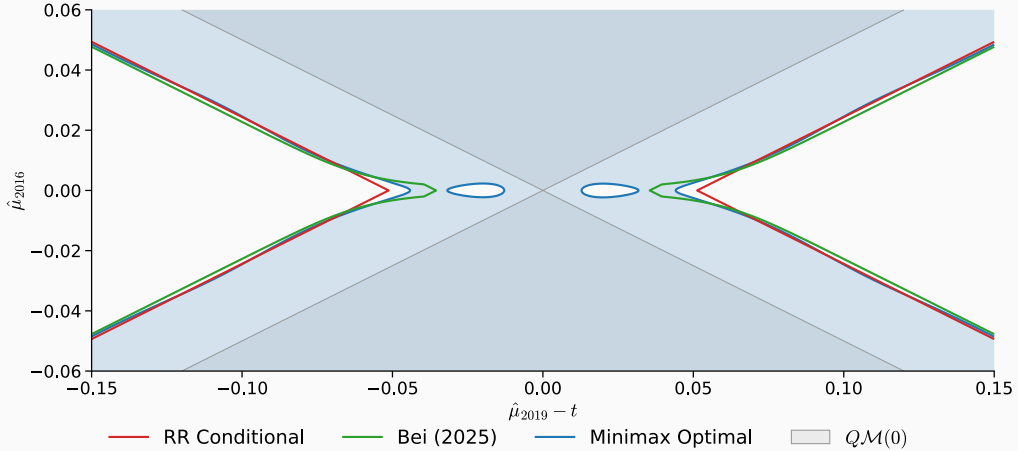
- Authors conduct sensitivity analysis on $\theta = \tau_{2019}$ under $\Delta^{\text{RM}}(\bar{M})$

Definitions

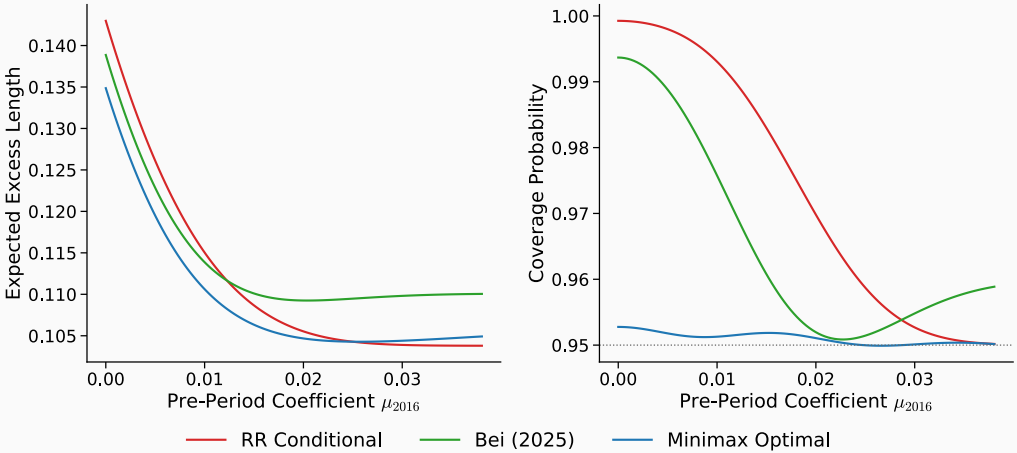
Iovino et al. (2026) under $\Delta^{\mathrm{RM}}(\bar{M})$



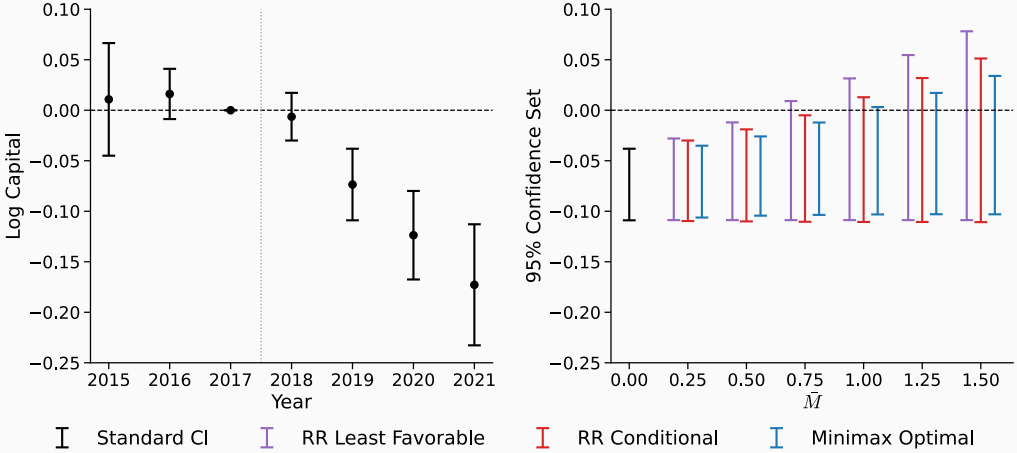
Acceptance Regions for $\Delta^{\mathsf{RM}}(1)$, Single Pre-Period



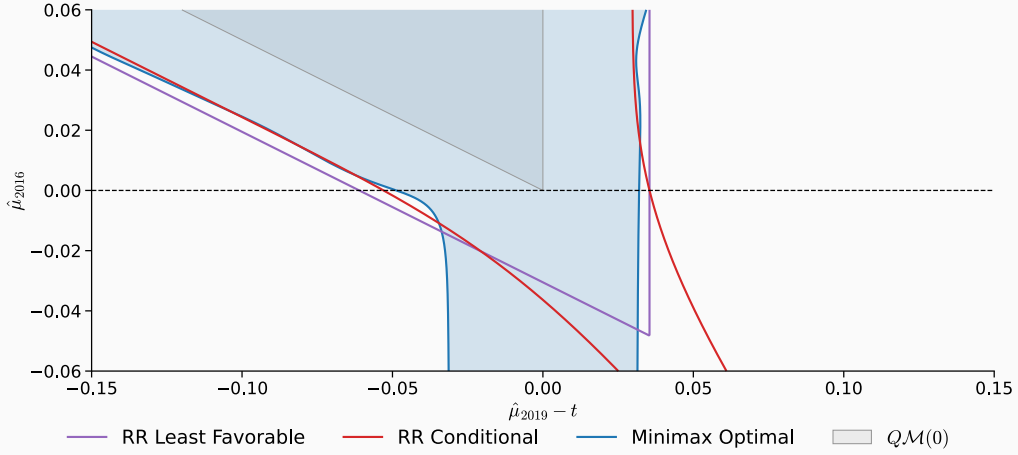
Risk and Coverage for $\Delta^{\mathsf{RM}}(1)$, Single Pre-Period



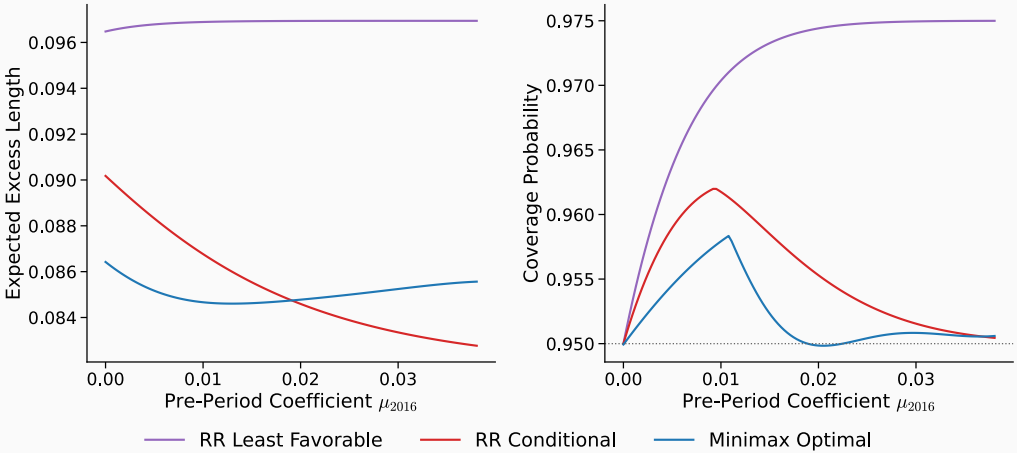
Iovino et al. (2026) under $\Delta^{\mathsf{RMD}}(\bar{M})$



Acceptance Regions for $\Delta^{\mathsf{RMD}}(1)$, Single Pre-Period



Risk and Coverage for $\Delta^{\mathsf{RMD}}(1)$, Single Pre-Period



Conclusions

- Little is known about the optimality properties of confidence procedures for partially identified parameters
- I propose a precise notion of optimality by taking a decision-theoretic perspective
- I establish that for a flexible class of models, inverting a single optimal test yields a minimax optimal confidence procedure
- I approximate the optimal procedure and quantify the suboptimality of existing methods

Definition of the Action Space

Formally, I take the action space to be the measure algebra: the quotient of the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}$ by the σ-ideal of null sets,

$$\mathcal{A} \equiv \mathcal{B}/\{A \in \mathcal{B} : \lambda(A) = 0\}.$$

Unlike $\mathcal{B}$, the measure algebra is endowed with a natural topology under which it is a Polish space (Le Maître, 2026). I equip $\mathcal{A}$ with the corresponding Borel σ-algebra.

Return

Definition of the Coverage Requirement

Problem: The indicator $\mathbb{1}\{t \in S(X,U)\}$ is not well-defined.

Solution: Let $I_S : \mathbb{R} \times \mathbb{R}^{k_\mu} \times [0,1] \to [0,1]$ be a measurable map such that for every (x,u), the slice $I_S(\cdot\,, x, u)$ is a version of $\mathrm{d}\lambda_{S(x,u)}/\mathrm{d}\lambda$. Writing $C(t,\mu,S) \equiv \mathbb{E}_{P_{\mu,U}}[I_S(t,X,U)]$ for the coverage probability function,

$$\mathcal{S}_\alpha \equiv \Big\{ S \in \mathcal{S} : \operatorname*{ess\,inf}_{t\in\mathbb{R}} \inf_{\mu \in \mathcal{M}(t)} C(t,\mu,S) \geq 1-\alpha \Big\}.$$

Return

Definition of $\Delta^{\text{SDRM}}(\bar{M})$

Second Differences Relative Magnitudes

$$\Delta^{\text{SDRM}}(\bar{M}) \equiv \Big\{ \delta : |(\delta_{t+1} - \delta_t) - (\delta_t - \delta_{t-1})| \leq \bar{M} \max_{s<0} |(\delta_{s+1} - \delta_s) - (\delta_s - \delta_{s-1})| \text{ for all } t \geq 0 \Big\}$$

Return

The Neyman–Pearson Test

Let p_ν denote the density of P_ν, and for a measure w on $\mathbb{R}^{k_\nu}$ define the mixture of densities

$$p_w(y) \equiv \int p_\nu(y)\, w(\mathrm{d}\nu)$$

A prior $\pi \in \Pi$ induces weights over the alternative,

$$\omega_\pi(\mathrm{d}\nu) \equiv \mathbb{1}\{\nu \notin Q\mathcal{M}(0)\}\, \mathrm{d}\nu_1\, \pi(\mathrm{d}\nu_{-1})$$

Neyman–Pearson Test

For $\pi \in \Pi$ and $\Gamma \in \mathcal{G}$,

$$\varphi_{\pi,\Gamma}(y) \equiv \mathbb{1}\{p_{\omega_\pi}(y) > p_\Gamma(y)\}$$

Return

Stochastic Mirror Ascent on $(\tilde{\pi}, \tilde{\Gamma})$

Pseudo-Code for Solving the Discretized Dual Problem

input: grids on $Q_{-1}\mathcal{M}$ and $Q\mathcal{M}(0)$, step sizes $\eta_L, \eta_M > 0$, iteration count T

initialize: $\tilde{\pi}$ and $\tilde{\Gamma}$ uniform on their grids

repeat T times:

1. $\varphi \leftarrow \mathbb{1}\{p_{\omega_{\tilde{\pi}}} > p_{\tilde{\Gamma}}\}$
2. Monte Carlo estimates $\hat{R}(\nu_{-1})$ of $R(\nu_{-1}, \varphi)$ and $\hat{\alpha}(\nu)$ of $\mathbb{E}_{P_\nu}[\varphi(Y)]$ on the grids
3. $\tilde{\pi}(\nu_{-1}) \leftarrow \tilde{\pi}(\nu_{-1}) \exp(\eta_L \hat{R}(\nu_{-1}))$ for each ν_{-1}, then renormalize
4. $\tilde{\Gamma}(\nu) \leftarrow \tilde{\Gamma}(\nu) \exp(\eta_M(\hat{\alpha}(\nu) - \alpha))$ for each ν

return the averages of $\tilde{\pi}$ and $\tilde{\Gamma}$ across iterations, and the test φ they induce

Return

Andrews, I., J. Roth, and A. Pakes (2023). Inference for linear conditional moment inequalities. *Review of Economic Studies 90*(6), 2763–2791.

Aradillas Fernández, A., J. Blanchet, J. L. Montiel Olea, C. Qiu, J. Stoye, and L. Tan (2025, November). Approximate least-favorable distributions and nearly optimal tests via stochastic mirror descent. arXiv preprint.

Armstrong, T. B. and M. Kolesár (2018). Optimal inference in a class of regression models. *Econometrica 86*(2), 655–683.

Armstrong, T. B. and M. Kolesár (2021). Sensitivity analysis using approximate moment condition models. *Quantitative Economics 12*(1), 77–108.

Bei, X. (2025, October). Inference on union bounds. Working paper, The University of Texas at Austin.

Canay, I. A., G. Illanes, and A. Velez (2023). A user's guide for inference in models defined by moment inequalities. *Journal of Econometrics*, 105558.

Cox, G. F. (2026). A simple and adaptive confidence interval when nuisance parameters satisfy an inequality. *Journal of the American Statistical Association* (just-accepted), 1–51.

de Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and bounding experimenter demand. *American Economic Review 108*(11), 3266–3302.

Donoho, D. L. (1994). Statistical estimation and optimal recovery. *The Annals of Statistics 22*(1), 238–270.

Ghosh, J. K. (1961). On the relation among shortest confidence intervals of different types. *Calcutta Statist. Assoc. Bull. 10*(4), 147–152.

Gundersen, C., B. Kreider, and J. Pepper (2012). The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics 166*(1), 79–91.

Imbens, G. W. and C. F. Manski (2004). Confidence intervals for partially identified parameters. *Econometrica 72*(6), 1845–1857.

Iovino, L., T. Martin, and J. Sauvagnat (2026). The environmental bias of corporate income taxation. *American Economic Review 116*(9), 3510–3551.

Katz, M. (2007). *Supermarkets and Zoning Laws*. Ph. D. thesis, Harvard University.

Ketz, P. and A. McCloskey (2025). Short and simple confidence intervals when the directions of some effects are known. *The Review of Economics and Statistics 107*(3), 820–834.

Kolesár, M., J. L. Montiel Olea, and J. Roth (2026, March). Evaluating counterfactual policies using instruments. arXiv preprint.

Kolesár, M. and C. Rothe (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review 108*(8), 2277–2304.

Krafft, O. and H. Witting (1967). Optimale Tests und ungünstigste Verteilungen. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 7*, 289–302.

Le Maître, F. (2026). On dense orbits in the space of subequivalence relations. *Orbita Mathematicae 3*(1), 59–87.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies 76*(3), 1071–1102.

Li, C. and U. K. Müller (2021). Linear regression with many controls of limited explanatory power. *Quantitative Economics 12*(2), 405–442.

Manski, C. F. (1990). Nonparametric bounds on treatment effects. *The American Economic Review 80*(2), 319–323.

Manski, C. F. and J. V. Pepper (2018). How do right-to-carry laws affect crime rates? coping with ambiguity using bounded-variation assumptions. *Review of Economics and Statistics 100*(2), 232–244.

Masten, M. A. and A. Poirier (2021). Salvaging falsified instrumental variable models. *Econometrica 89*(3), 1449–1469.

Nevo, A. and A. M. Rosen (2012). Identification with imperfect instruments. *Review of Economics and Statistics 94*(3), 659–671.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2015). Moment inequalities and their application. *Econometrica 83*(1), 315–334.

Pratt, J. W. (1961). Length of confidence intervals. *Journal of the American Statistical Association 56*(295), 549–567.

Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies 90*(5), 2555–2591.

Stein, C. M. (1962). Confidence sets for the mean of a multivariate normal distribution. *Journal of the Royal Statistical Society. Series B (Methodological) 24*(2), 265–296.

Wald, A. (1950). *Statistical Decision Functions*. New York: John Wiley & Sons.