A Simple, Never-Empty Confidence Interval for Partially Identified Parameters

Jörg Stoye

EEA-ESEM Summer Meeting, Milan, 2022

A Simple, Never-Empty Confidence Interval

This paper was triggered by an inquiry from the authors of (AER, 2018):

Measuring and Bounding Experimenter Demand[†]

By Jonathan de Quidt, Johannes Haushofer, and Christopher Roth*

We propose a technique for assessing robustness to demand effects of findings from experiments and surveys. The core idea is that by deliberately inducing demand in a structured way we can bound its influence. We present a model in which participants respond to their beliefs

The essence of their problem:

- Estimating upper and lower bounds (θ_L, θ_U) on some effect θ.
- The bounds come from separate subsamples.
- As a result, occasionally $\hat{\theta}_L > \hat{\theta}_U$.
- Established CI $[\hat{\theta}_L - 1.96 SE_L, \hat{\theta}_U + 1.96 SE_U]$ can be short or empty.
- Isn't such inference spuriously precise?

That is a very good question!

It has been discussed in econometrics:

- In the refereeing process of my own 2009 paper.
- Ponomareva and Tamer (2011) first to point out spurious precision.
- Molinari (2020) discusses in detail.
- Andrews and Kwon (2019):
  General but delicate (several tuning parameters) treatment.
- This paper is the opposite extreme:
  Only the motivating example.

A Simple, Never-Empty Confidence Interval

The Setting

We are interested in $\theta \in [\theta_L, \theta_U]$.

We have estimators

$$\begin{pmatrix} \hat{\theta}_L \\ \hat{\theta}_U \end{pmatrix} \sim N\left(\begin{pmatrix} \theta_L \\ \theta_U \end{pmatrix}, \begin{pmatrix} \sigma_L^2 & \rho\sigma_L\sigma_U \\ \rho\sigma_L\sigma_U & \sigma_U^2 \end{pmatrix} \right)$$

with $(\sigma_L, \sigma_U, \rho)$ known.

- Of course, this can be the *asymptotic* experiment. Still restrictive.
- Counterexample: Intersection bounds.
- Of special interest: $\rho = 0$ (independent samples).

A Simple, Never-Empty Confidence Interval

Following Imbens and Manski (2004), we want a confidence interval for θ:

$$\inf_{\theta \in [\theta_L, \theta_U]} \Pr(\theta \in CI) = .95.$$

The easiest way to do this depends on $\Delta \equiv \theta_U - \theta_L$.

If Δ is large, inference is one-sided and we can use

$$CI_\infty \equiv [\hat{\theta}_L - 1.64\sigma_L, \hat{\theta}_U + 1.64\sigma_U].$$

But if $\Delta \approx 0$, we need to use

$$CI_0 \equiv [\hat{\theta}_L - 1.96\sigma_L, \hat{\theta}_U + 1.96\sigma_U].$$

(Can be minimally refined using ρ).

A Simple, Never-Empty Confidence Interval

Following Imbens and Manski (2004), we want a confidence interval for θ:

$$\inf_{\theta \in [\theta_L, \theta_U]} \Pr(\theta \in CI) = .95.$$

The easiest way to do this depends on $\Delta \equiv \theta_U - \theta_L$.

If Δ is large, inference is one-sided and we can use

$$CI_\infty \equiv [\hat{\theta}_L - 1.64\sigma_L, \hat{\theta}_U + 1.64\sigma_U].$$

But if $\Delta \approx 0$, we need to use

$$CI_0 \equiv [\hat{\theta}_L - 1.96\sigma_L, \hat{\theta}_U + 1.96\sigma_U].$$

(Can be minimally refined using ρ).

Naively basing the case distinction on $\hat{\Delta} = \hat{\theta}_U - \hat{\theta}_L$ will fail if $\Delta \approx 0$.

This is well understood and is related to *post-model selection inference* as well as *inference near boundary of parameter space*.

A Simple, Never-Empty Confidence Interval

The fix:
Conservative pre-test or shrinkage estimation
(Stoye, 2009; see also Andrews and Soares, 2010; Bugni, 2010; Canay, 2010).

We shrink $\hat{\Delta}$ to 0 to ensure $\Delta \approx 0 \Rightarrow \hat{\Delta} = 0$ (in some stochastic sense).

A separate problem:
The interval can be empty.

A Simple, Never-Empty Confidence Interval

What if we don't want that?

Idea (Andrews/Kwon 2019):
Force coverage of pseudotrue parameter value

$$\theta^* = \frac{\sigma_L \theta_U + \sigma_U \theta_L}{\sigma_L + \sigma_U}.$$

Henceforth, a CI is valid if

$$\inf_{\theta \in \{\theta^*\} \cup [\theta_L, \theta_U]} \Pr(\theta \in CI) \geq 1 - \alpha.$$

A Simple, Never-Empty Confidence Interval

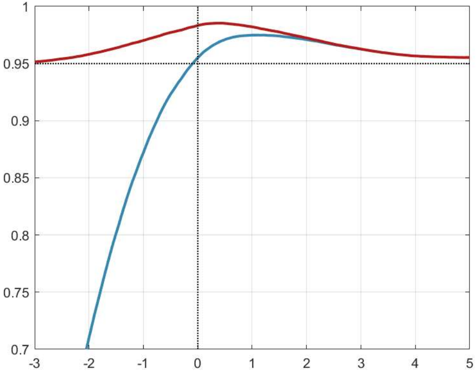
An obvious CI for θ^*:

$$\begin{aligned} CI_{\theta^*} &\equiv \left[\hat{\theta}^* - 1.96\sigma^*, \hat{\theta}^* - 1.96\sigma^*\right] \\ \hat{\theta}^* &\equiv \frac{\sigma_L \hat{\theta}_U + \sigma_U \hat{\theta}_L}{\sigma_L + \sigma_U} \\ \sigma^* &\equiv \frac{\sqrt{2+2\rho}\sigma_L \sigma_U}{\sigma_L + \sigma_U}. \end{aligned}$$

A very heavy-handed fix:

Depending on pre-test, report $CI_0 \cup CI_{\theta^*}$ or $CI_\infty \cup CI_{\theta^*}$.

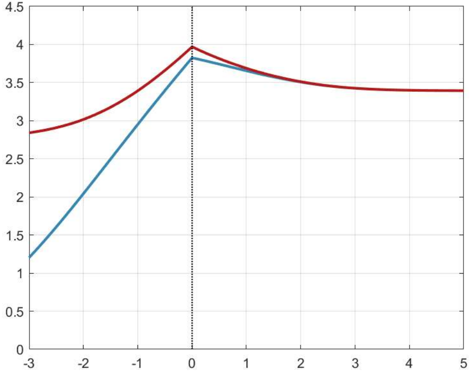
A Simple, Never-Empty Confidence Interval



- Horizontal axis is $\Delta \equiv \theta_U - \theta_L$, the length of Θ_I.
- $\Delta = 0$ is point identification. $\Delta < 0$ is misspecification.

We see **coverage** for traditional CI and heavy-handed fix with $\rho = 0$.

A Simple, Never-Empty Confidence Interval



- Horizontal axis is $\Delta \equiv \theta_U - \theta_L$, the length of Θ_I.
- $\Delta = 0$ is point identification. $\Delta < 0$ is misspecification.

We see **length** for traditional CI and heavy-handed fix with $\rho = 0$.

A Simple, Never-Empty Confidence Interval

New proposal:

- Take CI_{θ^*} into account when calibrating critical value.
- Problem: Effect of Δ on coverage is not any more monotonic.
- In higher dimension, this makes for a very hard problem.
- We will concentrate out globally least favorable value of $\Delta \geq 0$.
- No shrinkage, pre-testing, or tuning parameter.

A Simple, Never-Empty Confidence Interval

New proposal:

- Take $CI_{\theta*}$ into account when calibrating critical value.
- Problem: Effect of Δ on coverage is not any more monotonic.
- In higher dimension, this makes for a very hard problem.
- We will concentrate out globally least favorable value of $\Delta \geq 0$.
- No shrinkage, pre-testing, or tuning parameter.

This is pretty straightforward. But now things get interesting.

- For a wide range of cases, $\Delta = \infty$ is globally least favorable.
- In those cases, can just report $CI_\infty \cup CI_{\theta*}$.
- This is provably the case if $\rho = 0$.
- In that case:
  Just use a critical value of 1.64 (plus union with $CI_{\theta*}$) to get .95 coverage.
- Not obvious! (Discovered by simulation. I initially assumed a bug.)

Formal Statement

Report

$$CI_{MA} \equiv CI_{\theta^*} \cup [\hat{\theta}_L - c\sigma_L, \hat{\theta}_U + c\sigma_U],$$

where c is calibrated such that

$$\inf_{\theta \in \{\theta^*\} \cup [\theta_L, \theta_U]} \Pr(\theta \in CI_{MA}) \geq .95$$

irrespective of the value taken by Δ.

Formally, after some simplification justified in the paper:

$$\inf_{\Delta \geq 0} \Pr\left(Z_1 - \Delta - c \leq 0 \leq Z_2 + c \text{ or } |Z_1 + Z_2 - \Delta| \leq \sqrt{2 + 2\rho} \cdot 1.96\right) = .95,$$

$$\text{where } \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

A Simple, Never-Empty Confidence Interval

Formal Statement

Report

$$CI_{MA} \equiv CI_{\theta^*} \cup [\hat{\theta}_L - c\sigma_L, \hat{\theta}_U + c\sigma_U],$$

where c is calibrated such that

$$\inf_{\theta \in \{\theta^*\} \cup [\theta_L, \theta_U]} \Pr(\theta \in CI_{MA}) \geq .95$$

irrespective of the value taken by Δ.

Why concentrate out Δ rather than using an estimator $\hat{\Delta}$?

- As before, one would then have to adjust for pre-estimation. Among other things, this would introduce a tuning parameter.
- The coverage probability is not monotonic in Δ, so shrinking Δ in a specific direction will not work.
- Can prove that it suffices to consider coverage only at θ_U, so computation is an easy grid search on the real line.

Formal Statement

Report

$$CI_{MA} \equiv CI_{\theta^*} \cup [\hat{\theta}_L - c\sigma_L, \hat{\theta}_U + c\sigma_U],$$

where c is calibrated such that

$$\inf_{\theta \in \{\theta^*\} \cup [\theta_L, \theta_U]} \Pr(\theta \in CI_{MA}) \geq .95$$

irrespective of the value taken by Δ.

Can easily compute c as a function of ρ.

$\boldsymbol{\rho}$	≤ 0.8	0.85	0.9	0.95	0.98	1
$\boldsymbol{\alpha = .1}$	1.28	1.29	1.31	1.36	1.44	1.64
$\boldsymbol{\alpha = .05}$	1.64	1.65	1.65	1.70	1.76	1.96
$\boldsymbol{\alpha = .01}$	2.33	2.33	2.33	2.34	2.40	2.58

A Simple, Never-Empty Confidence Interval

Formal Statement

Report

$$CI_{MA} \equiv CI_{\theta^*} \cup [\hat{\theta}_L - c\sigma_L, \hat{\theta}_U + c\sigma_U],$$

where c is calibrated such that

$$\inf_{\theta \in \{\theta^*\} \cup [\theta_L, \theta_U]} \Pr(\theta \in CI_{MA}) \geq .95$$

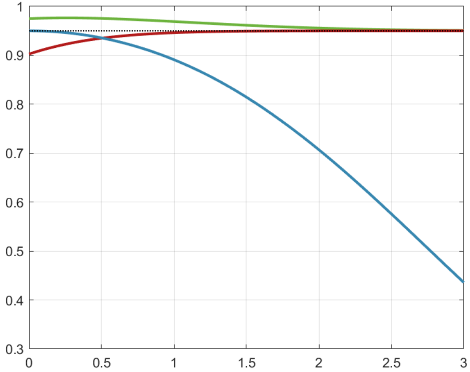
irrespective of the value taken by Δ.

Can easily compute c as a function of ρ.

$\boldsymbol{\rho}$	≤ 0.8	0.85	0.9	0.95	0.98	1
$\boldsymbol{\alpha = .1}$	1.28	1.29	1.31	1.36	1.44	1.64
$\boldsymbol{\alpha = .05}$	1.64	1.65	1.65	1.70	1.76	1.96
$\boldsymbol{\alpha = .01}$	2.33	2.33	2.33	2.34	2.40	2.58

Except for large ρ, it's just the one-sided critical value!
Again, for $\rho = 0$ this is provable.

A Simple, Never-Empty Confidence Interval



This plots coverage at $c = 1.64$ as function of $\Delta \geq 0$ for $\rho = 0$.

- Coverage is not monotonic in Δ.
- Coverage is minimized as $\Delta \to \infty$.

Bet Proofness

The new interval is bet proof.

What does that mean?

- Consider an “inspector" who can place a bet against coverage at nominally fair odds...
- ...after seeing the data (but before seeing the true parameter value).
- Bet-proofness obtains if there exists a parameter value for which the inspector **cannot** win on average.
- In nontrivial settings, this is hard to attain or prove (Mueller-Norets 2016, statistics literature before that).
- No interval that can be empty fulfils it:
  The inspector can bet against coverage for empty intervals.

A Simple, Never-Empty Confidence Interval

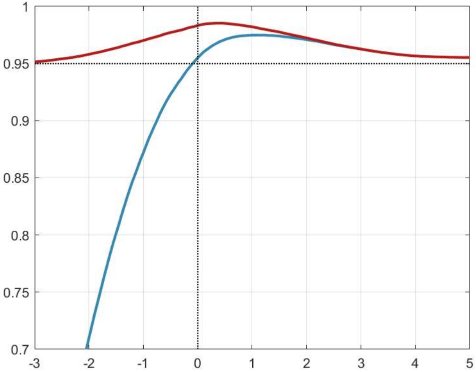
Bet Proofness

The new interval is bet proof.

Why is it true?

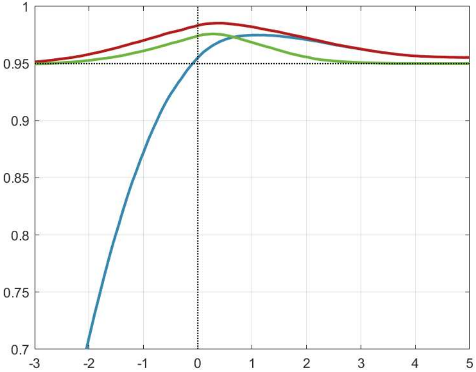
- The lucky parameter configuration is $\theta_L = \theta_U \Leftrightarrow \Theta_I = \{\theta_L\}$.
- In this case, we also have $\Theta_I = \{\theta^*\}$.
- By a theorem in Wallace (1959), CI_{θ^*} is bet-proof for θ^*.
- Bet-proofness extends to supersets.

A Simple, Never-Empty Confidence Interval



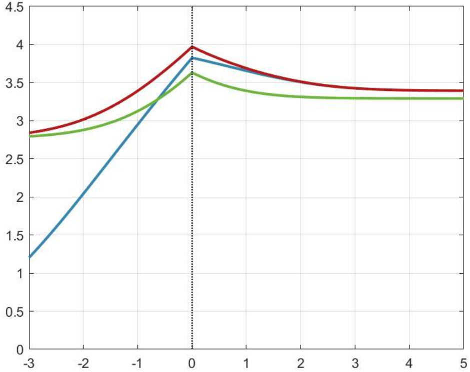
So how does the interval perform?
This is the previous graph.

A Simple, Never-Empty Confidence Interval



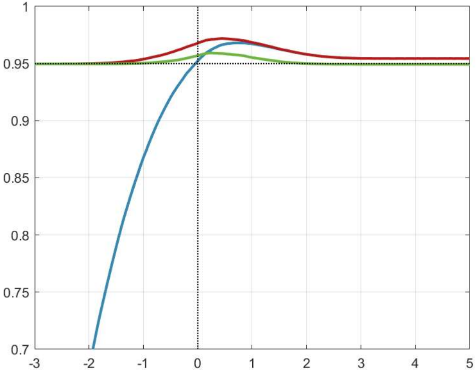
So how does the interval perform?
The green curve is the new interval's size control.

A Simple, Never-Empty Confidence Interval



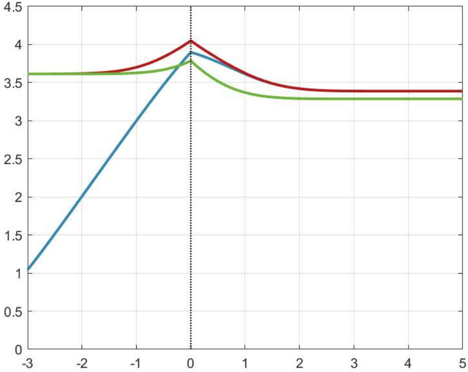
Here is the same comparison for length.

A Simple, Never-Empty Confidence Interval



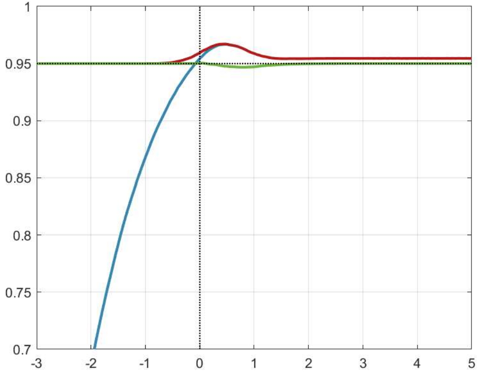
Here is coverage again, but now $\rho = .7$ ($\approx$ best case).

A Simple, Never-Empty Confidence Interval

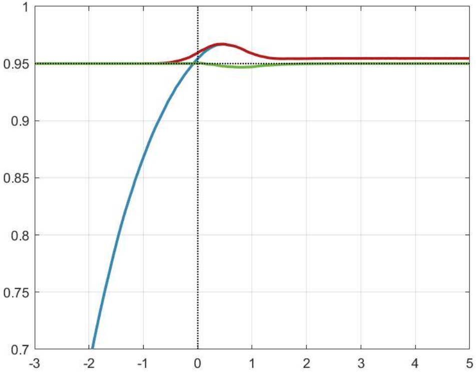


Here is length again, but now $\rho = .7$.

A Simple, Never-Empty Confidence Interval



So how does the interval perform?
Here is coverage again, but now using $c = 1.64$ at $\rho = .95$.



So how does the interval perform?

For very high ρ, worst-case Δ is interior and one-sided critical value does not do.

(And is not suggested; the figure is strictly illustrative.)

Empirical Application

Let's go back to the paper that started it all. For some experiments:

Game	$[\hat{\theta}_L, \hat{\theta}_U]$	CI_{MA}	CI_{TI}	rel. length
Ambiguity Aversion	[0.499,0.557]	[0.459,0.597]	[0.458,0.598]	0.97
Effort: 1 cent bonus	[0.469,0.484]	[0.448,0.503]	[0.448,0.504]	0.97
Effort: 0 cent bonus	[0.343,0.331]	[0.318,0.356]	[0.315,0.358]	0.91
Lying	[0.530,0.537]	[0.512,0.556]	[0.508,0.560]	0.83
Time	[0.766,0.770]	[0.722,0.814]	[0.712,0.824]	0.82
Trust Game 1	[0.430,0.455]	[0.388,0.493]	[0.387,0.495]	0.96
Trust Game 2	[0.348,0.398]	[0.328,0.426]	[0.327,0.427]	0.97
Ultimatum Game 1	[0.443,0.470]	[0.422,0.493]	[0.422,0.494]	0.97
Ultimatum Game 2	[0.362,0.413]	[0.342,0.436]	[0.341,0.436]	0.97

The new interval makes a difference.

Empirical Application

Let's go back to the paper that started it all. For some experiments:

Game	$[\hat{\theta}_L, \hat{\theta}_U]$	CI_{MA}	CI_{TI}	rel. length
Ambiguity Aversion	[0.499,0.557]	[0.459,0.597]	[0.458,0.598]	0.97
Effort: 1 cent bonus	[0.469,0.484]	[0.448,0.503]	[0.448,0.504]	0.97
Effort: 0 cent bonus	[0.343,0.331]	[0.318,0.356]	[0.315,0.358]	0.91
Lying	[0.530,0.537]	[0.512,0.556]	[0.508,0.560]	0.83
Time	[0.766,0.770]	[0.722,0.814]	[0.712,0.824]	0.82
Trust Game 1	[0.430,0.455]	[0.388,0.493]	[0.387,0.495]	0.96
Trust Game 2	[0.348,0.398]	[0.328,0.426]	[0.327,0.427]	0.97
Ultimatum Game 1	[0.443,0.470]	[0.422,0.493]	[0.422,0.494]	0.97
Ultimatum Game 2	[0.362,0.413]	[0.342,0.436]	[0.341,0.436]	0.97

The new interval makes a difference.

In one row, $\hat{\theta}_L > \hat{\theta}_U$. The new interval is in fact CI_{θ^*} and is still shorter.

Summary

For the simple but empirically salient case of interval identification with (asymptotically) jointly normal estimators, the new CI:

- Is never empty or very short.
- Is bet-proof in the asymptotic experiment.
- Covers a well-defined pseudotrue parameter even if the model is misspecified.
- Is trivial to compute and has no tuning parameters.
- Has commanding size control and length.
- Improves inference in the motivating empirical example.
- Has already seen another empirical application (Lee and Weidner, 2021).

Summary

For the simple but empirically salient case of interval identification with (asymptotically) jointly normal estimators, the new CI:

- Is never empty or very short.
- Is bet-proof in the asymptotic experiment.
- Covers a well-defined pseudotrue parameter even if the model is misspecified.
- Is trivial to compute and has no tuning parameters.
- Has commanding size control and length.
- Improves inference in the motivating empirical example.
- Has already seen another empirical application (Lee and Weidner, 2021).

Thank You!